Population genomics of the Viking world

Ashot Margaryan[1,2,3*], Daniel J. Lawson[4*], Martin Sikora[1*], Fernando Racimo[1*], Simon Rasmussen[5], Ida Moltke[6], Lara Cassidy[7], Emil Jørsboe[6,8], Andrés Ingason[1,9,10], Mikkel W. Pedersen[1], Thorfinn Korneliussen[1,11], Helene Wilhelmson[12,13], Magdalena M. Buś[14], Peter de Barros Damgaard[1], Rui Martiniano[15], Gabriel Renaud[1,34], Claude Bhérer[16], J. Víctor Moreno-Mayar[1,17], Anna K. Fotakis[3], Marie Allen[14], Raili Allmäe[18], Martyna Molak[19], Enrico Cappellini[3], Gabriele Scorrano[3], Hugh McColl[1], Alexandra Buzhilova[20], Allison Fox[21], Anders Albrechtsen[6], Berit Schütz[22], Birgitte Skar[23], Caroline Arcini[24], Ceri Falys[25], Charlotte Hedenstierna Jonson[26], Dariusz Błaszczyk[27], Denis Pezhemsky[20], Gordon Turner-Walker[28], Hildur Gestsdóttir[29], Inge Lundstrøm[3], Ingrid Gustin[12], Ingrid Mainland[30], Inna Potekhina[31], Italo M. Muntoni[32], Jade Cheng[1], Jesper Stenderup[1], Jilong Ma[1], Julie Gibson[30], Jüri Peets[18], Jörgen Gustafsson[33], Katrine H. Iversen[5,34], Linzi Simpson[35], Lisa Strand[23], Louise Loe[36], Maeve Sikora[37], Marek Florek[38], Maria Vretemark[39], Mark Redknap[40], Monika Bajka[41], Tamara Pushkina[42], Morten Søvsø[43], Natalia Grigoreva[44], Tom Christensen[45], Ole Kastholm[46], Otto Uldum[47], Pasquale Favia[48], Per Holck[49], Sabine Sten[50], Símun V. Arge[51], Sturla Ellingvåg[1], Vayacheslav Moiseyev[52], Wiesław Bogdanowicz[19], Yvonne Magnusson[53], Ludovic Orlando[54], Peter Pentz[45], Mads Dengsø Jessen[45], Anne Pedersen[45], Mark Collard[55], Daniel G. Bradley[7], Marie Louise Jørkov[56], Jette Arneborg[45,57], Niels Lynnerup[56], Neil Price[26], M. Thomas P. Gilbert[3,58], Morten E. Allentoft[1,59], Jan Bill[60], Søren M. Sindbæk[61], Lotte Hedeager[62], Kristian Kristiansen[63], Rasmus Nielsen[1,64†], Thomas Werge[1,9,10,65†], and Eske Willerslev[1,66,67,68†]

[1]Lundbeck Foundation GeoGenetics Centre, GLOBE Institute, University of Copenhagen, Øster Voldgade 5-7, 1350 Copenhagen, Denmark.
[2]Institute of Molecular Biology, National Academy of Sciences, 7, Hasratian St., 0014, Yerevan, Armenia.
[3]Section for Evolutionary Genomics, GLOBE Institute, University of Copenhagen, Øster Voldgade 5-7, 1350 Copenhagen, Denmark.
[4]MRC Integrative Epidemiology Unit and School of Statistical Sciences, University of Bristol, Bristol, UK.
[5]Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, 2200 Copenhagen, Denmark.
[6]Department of Biology, The Bioinformatics Centre, University of Copenhagen, 2200 Copenhagen, Denmark.
[7]Smurfit Institute of Genetics, Trinity College Dublin, Dublin.
[8]Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, 2200 Copenhagen, Denmark
[9]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark.
[10]Institute of Biological Psychiatry, Mental Health Services Copenhagen, Copenhagen, Denmark.
[11]HSE University, Russian Federation National Research University Higher School of Economics, 20 Myasnitskaya ulitsa, Moscow 101000 Russia.
[12]Historical archaeology, Department of Archaeology and Ancient history, Lund University, PB 192, SE 22100 Lund, Sweden.
[13]Sydsvensk arkeologi AB, PB 134, SE 29122 Kristianstad, Sweden.
[14]Department of Immunology, Genetics and Pathology, Uppsala University, 751 08 Uppsala, Sweden.
[15]Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK.

[16]Department of Human Genetics, McGill University, Montréal, Québec Canada.
[17]National Institute of Genomic Medicine (INMEGEN), Periférico Sur 4809, 14610 Mexico City, Mexico.
[18]Archaeological Research Collection, Tallinn University, Rüütli 10, Tallinn 10130, Estonia.
[19]Museum and Institute of Zoology, Polish Academy of Sciences, Wilcza 64, 00-679 Warsaw, Poland.
[20]Anuchin Research Institute and Museum of Anthropology, Moscow State University, Mokhovaya str.11, Moscow 125009, Russia.
[21]Manx National Heritage, Kingswood Grove, Douglas, Isle of Man, British Isles IM1 3LY.
[22]Upplandsmuseet, Drottninggatan 7, 753 10 Uppsala, Sweden.
[23]NTNU University Museum, Department of Archaeology and Cultural History Norway.
[24]The Archaeologists, National Historical Museums.
[25]Thames Valley Archaeological Services (TVAS), Reading, UK.
[26]Department of Archaeology and Ancient History, Uppsala University, Box 626, 751 26 Uppsala, Sweden.
[27]Institute of Archaeology, University of Warsaw, ul. Krakowskie Przedmieście 26/28, 00-927 Warsaw, Poland.
[28]Department of Cultural Heritage Conservation, National Yunlin University of Science and Technology, Douliou, Taiwan.
[29]Institute of Archaeology, Iceland. Bárugata 3, 101 Reykjavík, Iceland.
[30]UHI Archaeology Institute, University of the Highlands and Islands, Orkney College, Kirkwall, Orkney, KW15 1LX.
[31]Department of Bioarchaeology, Institute of Archaeology of National Academy of Sciences of Ukraine, 12 Geroiv Stalingrada Ave. 04210 Kyiv, Ukraine.
[32]Soprintendenza Archeologia, Belle Arti e Paesaggio per le Province di Barletta - Andria - Trani e Foggia, Via Alberto Alvarez Valentini, 8 - 71121 Foggia, Italy.
[33]Jönköping county museum, Jönköping, Sweden.
[34]Department of Health Technology, Section for Bioinformatics, Technical University of Denmark, DTU, 2800 Kgs. Lyngby, Denmark
[35]Trinity College Dublin.
[36]Heritage Burial Services, Oxford Archaeology, Janus House, Osney Mead, Oxford OX2 0ES, UK.
[37]National Museum of Ireland, Kildare Street, Dublin 2, Ireland.
[38]Institute of Archaeology, Maria Curie-Sklodowska University in Lublin, Pl. M. Curie-Sklodowska 4, 20-035 Lublin, Poland.
[39]Västergötlands museum, Box 253, 532 23 Skara Sweden.
[40]Department of History & Archaeology, Amgueddfa Cymru – National Museum Wales, Cathays Park, Cardiff, Wales, CF10 3NP.
[41]"Trzy Epoki" Archaeological Service, Poland.
[42]Historical faculty, Moscow State University, Lomonosovsky prospekt 27/4, Moscow 119192, Russia.
[43]Museum of Southwest Jutland.
[44]Department of Slavic-Finnish Archaeology, Institute for the History of Material Culture, Russian Academy of Sciences, Dvotsovaya Emb., 18, 191186, Saint-Petersburg, Russia.
[45]National Museum of Denmark, Frederiksholms Kanal 12, DK-1220 Copenhagen, Denmark.
[46]Roskilde Museum, Department of Research and Heritage, Sankt Ols Stræde 3, DK-4000 Roskilde, Denmark.
[47]Langelands Museum, Jens Winthersvej 12. 5900 Rudkøbing, Langeland, Denmark.
[48]Department of Humanities, University of Foggia, Via Arpi, 176, 71121 Foggia, Italy.

[49]Department of Molecular Medicine, Faculty of Medicine, University of Oslo.
[50]Department of Archaeology and Ancient History, Uppsala University Campus Gotland.
[51]Tjóðsavnið - Faroe Islands National Museum. Kúrdalsvegur 15. Postboks 1155. FO-110 Tórshavn.
[52]Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), Russian Academy of Science, University Emb, 3, SPb, Russia, 199034.
[53]Malmö Museum, Box 406, 201 24 Malmö, Sweden.
[54]Laboratoire d'Anthropobiologie Moléculaire et d'Imagerie de Synthèse, CNRS UMR 5288, Université de Toulouse, Université Paul Sabatier, 31000 Toulouse, France.
[55]Department of Archaeology, Simon Fraser University, 8888 University Dr, Burnaby, BC V5A 1S6, Canada.
[56]Department of Forensic Medicine, University of Copenhagen, Frederik V's vej 11, 2100 Copenhagen.
[57]School of GeoSciences, University of Edinburgh.
[58]Department of Natural History, NTNU: Norwegian University of Science and Technology.
[59]Trace and Environmental DNA (TrEnD) Laboratory, School of Molecular and Life Sciences, Curtin University, Kent Street, 6102 Perth, Australia.
[60]Museum of Cultural History, University of Oslo, P.O. Box 6762 St. Olavs plass, 0160 Oslo, Norway.
[61]Centre for Urban Network Evolutions (UrbNet), Aarhus University, School of Culture and Society, Moesgård Allé 20, building 4215, DK-8270 Højbjerg, Denmark.
[62]Institute of Archaeology, Conservation and History, Pb. 1019 Blindern, 0315 Oslo, Norway.
[63]Department of Historical Studies, University of Gothenburg.
[64]Departments of Integrative Biology and Statistics, UC Berkeley, Berkeley, CA 94720, USA.
[65]The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Denmark.
[66]Department of Zoology, University of Cambridge, UK.
[67]The Danish Institute for Advanced Study, University of Southern Denmark.
[68]The Wellcome Trust Sanger Institute, Cambridge, UK.

*These authors contributed equally to this work.
[†]e-mail: ewillerslev@snm.ku.dk; Thomas.Werge@regionh.dk; rasmus_nielsen@berkeley.edu

Abstract

The Viking Age maritime expansion of Scandinavian populations (c. 750 to 1050 CE) was a far-flung transformation in world history[1,2]. To understand its global influence, we sequenced the genomes of 442 ancient humans (median depth of c. 1X) from across Europe and Greenland. We find the Viking period involved foreign gene flow into Scandinavia from the south and east. We observe genetic structure within Scandinavia, with diversity hotspots to the south and restricted gene flow within Scandinavia. We find evidence for a major Danish influx in England, Swedish influx in the Baltic, and Norwegian influx in Ireland, Iceland, and Greenland. Additionally, we see substantial foreign European ancestry entering Scandinavia during the Viking Age. We show that a Viking expedition included close family members. We find that pigmentation-associated loci have undergone strong population differentiation during the last millennia. We trace positively selected loci with unprecedented detail, including the lactase persistence allele and alleles associated with the immune response. We conclude that the Viking diaspora was characterized by substantial trans-regional engagement: distinct populations

influenced the genomic makeup of different regions of Europe, while Scandinavia experienced increased contact with the rest of the continent.

Introduction

The events of the Viking Age (VA) altered the political, cultural, and demographic map of Europe in ways that are evident to this day. Scandinavian diasporas established trade and settlement stretching from the American continent to the Asian steppe[1]. They exported ideas, technologies, language, beliefs, and practices to these lands, whilst developing new socio-political structures, and assimilating cultural influences[2].

To explore the genomic history of the VA, we "shotgun" sequenced DNA extracted from 442 ancient human remains dating from the Bronze Age (BA; c. 2400 BCE) to the Early Modern period (c. 1600 CE) (Fig. 1; Extended Data Fig. 1). The data from ancient individuals were analyzed together with published data from 3,855 present-day individuals across two reference panels (Supplementary Note 6), and data from 1,118 ancient individuals (Supplementary Table 3).

Scandinavian genetic ancestry and the beginnings of the Viking era

Although VA Scandinavians shared a common cultural background, there was no common word for Scandinavian identity at that time[1]. Rather than a single "Viking world", a series of interlinked "Viking worlds" emerged from rapidly growing maritime exploration, trade, war, and settlement, following the adoption of deep-sea navigation among coastal populations of Scandinavia and the Baltic Sea area[3,4]. Thus, it is unclear to what extent the Viking phenomenon refers to people with a recently shared genetic background or how far population changes accompanied the transition from the Iron Age (IA) to the VA in Scandinavia.

The VA Scandinavians of our study fall broadly within the diversity of ancient European individuals from the Bronze Age and later (Fig. 2; Extended Data Figs. 2 and 3; Supplementary Note 8), but with subtle differences among the different groups indicating complex fine-scale structure. For example, many VA individuals from the island of Gotland cluster with BA individuals from the Baltic region, indicating mobility across the Baltic Sea (Fig. 2 and Extended Data Fig. 3). Using f_4-statistics to contrast genetic affinities with Steppe pastoralists and Neolithic farmers, we find that VA individuals from Norway are distributed in a similar manner to earlier IA individuals, whereas many VA individuals from Sweden and Denmark show greater affinity to Neolithic farmers from Anatolia (Extended Data Fig. 4a). Using *qpAdm*, we find that the majority of groups can be modelled as three-way mixtures of hunter-gatherer, farmer, and Steppe-related ancestry. The three-way model was rejected for some groups from Sweden, Norway, and the Baltic region, which could be fit using a four-way model including Caucasus hunter-gatherer or East Asian-related ancestry (Extended Data Figs. 4b and 4c), the latter consistent with previously documented gene flow from Siberia[5–7]

Investigating genetic continuity between more temporally proximate IA groups and VA Scandinavians, we find that most VA groups can be fit using a single IA source, and broadly fall into two categories: i) English IA sources (most Danish VA, British Isles), and ii) Scandinavian IA sources (Norway, Sweden, and the Baltic) (Extended Data Fig. 5a). Notable exceptions are individuals from Kärda in Southern Sweden, for which only the early Medieval Longobard individuals from Hungary

can be fit as a single source group ($p > 0.01$; Extended Data Fig. 5a). Groups with poor one-way fits can be modelled by including either additional northeastern ancestry (e.g. Ladoga VA) or additional southeastern ancestry (e.g. Jutland VA) (Extended Data Fig. 5b). Overall, our analyses suggest that the genetic makeup of VA Scandinavians largely derives from ancestry of the preceding IA populations, but they also reveal subtle differences in ancestry and gene flow from both the south and east. These observations are largely consistent with archaeological findings[8,9].

Genetic structure within VA Scandinavia

To elucidate the fine-scale population structure of VA Scandinavia, we performed genotype imputation on a subset of 298 individuals with sufficient (>0.5X) coverage (289 from this study + 9 published[10]) and inferred genomic segments shared via identity-by-descent (IBD) with a reference panel of present-day Europeans (n=1,464, Supplementary Notes 6, 10 and 11). Genetic clustering using MDS and uniform manifold approximation and projection (UMAP) shows VA Scandinavians clustering into three groups by geographic origin, with close affinities to their respective present-day counterparts (Fig. 3a, Fig.S10.1). Some individuals have strong affinities with Eastern Europeans, particularly those from the island of Gotland in eastern Sweden, which likely reflects individuals with Baltic ancestry, as clustering with Baltic BA individuals is evident in the identity-by-state (IBS)-UMAP analysis (Fig. 2b) and through f_4-statistics (Fig S9.1).

We used ChromoPainter[11] and a reference panel enriched with Scandinavian individuals (n=1,464, see Supplementary Notes 6 and 11) to identify long, shared haplotypes and detect subtle population structure (Supplementary Figures S11.1-10). We find ancestry components in Scandinavia with (inexact and indicative) affinities with present-day populations (Fig. S11.11): “Danish-like”, “Swedish-like”, “Norwegian-like”, and “North Atlantic-like” (i.e. possibly individuals from the British Isles entering Scandinavia). The sampling is heavily structured, so these complex results (Fig. S11.12) are visualised over time and space (Fig. 4) using spatial interpolation[12] to account for sampling locations and report significant linear regressions (Supplementary Notes 11-12).

“Norwegian-like” and “Swedish-like” components cluster in Norway and Sweden, respectively, while “Danish-like” and “North Atlantic-like” components are widespread (Fig. 4, S11.12 and Supplementary Table 6). Unexpectedly, VA individuals from Jutland (Denmark) lack “Swedish-like” and “Norwegian-like” genetic components (Fig. S11.12). We also find that gene flow within Scandinavia was broadly from south to north, dominated by Danish movement into Norway and Sweden (Table S11.2).

We identified two ancient individuals from northern Norway (VK518, VK519) with affinities to present-day Saami in Norway and Sweden. The VK519 individual likely also had “Norwegian-like” ancestors, indicating genetic contacts between Saami and other Scandinavians populations.

The genetic data are structured by topographic boundaries rather than by present-day country borders. Thus, the south-western part of Sweden in the VA is genetically more similar to Danish VA populations than to central mainland Sweden, likely due to geographic barriers that prevented gene flow.

We quantified genetic diversity using two measures: conditional nucleotide diversity (Supplementary Note 9) and variation in inferred ancestry based on ChromoPainter results (Supplementary Note 11;

Extended Data Fig. 6 and Fig. S11.13). We also visualized it as the spread of individuals on the MDS plot based on a pairwise IBS sharing matrix (Fig. 3b).

Diversity varies significantly from more homogeneous inland and northern parts of Scandinavia to diverse Kattegat (eastern Denmark and western Sweden) and Baltic Sea regions, suggesting an important role for these maritime regions in interaction and trade during the VA. Interestingly, on Gotland, there are many more "Danish-like", "North Atlantic-like", and "Finnish-like" genetic components than "Swedish-like" components, indicating extensive maritime contacts during the VA.

Our results for Gotland and Öland agree with archaeological indications that these were important maritime communities from the Roman period onwards[13,14]. A similar pattern is observed on the central Danish islands, such as Langeland, but at a lower level. The data indicate that genetic diversity on the islands increased from early to late VA, suggesting increasing interregional interaction. Evidence for genetic structure within VA Scandinavia[2,4,15–17] with diversity in cosmopolitan centers like Skara and trade-oriented islands like Gotland, highlight the importance of sea routes.

Viking migrations

Our fine-scale ancestry analyses of genomic data are consistent with patterns documented by historians and archaeologists (Figs. 3, 4 and S11.12): eastward movements mainly involved "Swedish-like" ancestry, while individuals with "Norwegian-like" ancestry travelled to Iceland, Greenland, Ireland, and the Isle of Man. The first settlement in Iceland and Greenland also included individuals with "North Atlantic-like" ancestry[18,19]. A "Danish-like" ancestry is seen in present-day England, in accordance with historical records[20], place-names[21], surnames[22], and modern genetics[23,24], but VA "Danish-like" ancestry in the British Isles cannot be distinguished from that of the Angles and Saxons, who migrated in the 5th to 6th centuries CE from Jutland and Northern Germany.

VA execution sites in Dorset and Oxford, England, have significant "North Atlantic-like" ancestry as well as "Danish-like" and "Norwegian-like" ancestries. If these represent Viking raiding parties that were defeated and captured[25,26] then they were composed of individuals of different origins. This pattern is also suggested by isotopic data from a warrior cemetery in Trelleborg, Denmark[27]. Similarly, the presence of "Danish-like" ancestry in an ancient sample from Gnezdovo in present-day Russia indicates that eastern migrations were not entirely composed of Vikings from Sweden.

Importantly, our results show that "Viking" identity was not limited to individuals of Scandinavian genetic ancestry. Two Orkney individuals who were buried in Scandinavian fashion are genetically similar to present-day Irish and Scottish populations and are likely the first Pictish genomes published ("Evidence for Pictish Genomes", Supplementary Note 11, Figs S11.3, S11.12, S11.14, Supplementary Table 6). Two other Orkney individuals had 50% Scandinavian ancestry, and five such individuals were found in Scandinavia. This suggests that Pictish populations may have been integrated into Scandinavian culture by the VA.

Gene flow into Scandinavia during the Viking era

Non-Scandinavian ancestry in samples from Denmark, Norway, and Sweden agrees with known trading routes (Supplementary Notes 11 and 12). For example, Finnish and Baltic ancestry reached modern Sweden, including Gotland, but is absent in most individuals from Denmark and Norway. By contrast, western regions of Scandinavia received ancestry from the British Isles (Supplementary Notes 11 and 12). The first evidence of South European ancestry (>50%) in Scandinavia is during the VA in Denmark (e.g. VK365 and VK286 from Bogøvej) and southern Sweden (e.g. VK442 and VK350 from Öland, and VK265 from Kärda) (Fig. 4, Supplementary Table 6).

Disappearance of the Greenlandic Norse

From around 980 to 1440 CE southwest Greenland was settled by people of Scandinavian ancestry, probably from Iceland[28,29]. The fate of the Norse in Greenland remains debated, but probable causes of their disappearance are social or economic processes in Europe (e.g. political relations within Scandinavia and changed trading systems) and natural processes, including climatic change[29–31].

According to our data, the Greenlandic Norse were an admixture between Scandinavians (mostly from Norway) and individuals from the British Isles, similar to the first settlers of Iceland[18]. We see no evidence of long-term inbreeding in Greenlandic Norse genomes, though we have only one high-coverage genome from the later period of occupation of the island (Supplementary Note 10; Figs. S10.2 and S10.3). This result could favor a relatively brief depopulation scenario, in line with previous demographic models[32] and archaeological findings. We also find no evidence of ancestry from other populations (Paleo Eskimo, Inuit, or Native American) in the Greenlandic Norse genomes (Fig. S9.4), which accords with the skeletal remains[32]. This suggests that sexual interaction was absent or on a very small scale.

Genetic composition and kinship of the earliest Viking expedition

Whilst maritime raiding has been a constant of seafaring cultures for millennia, the VA is partly defined by this activity[33]. However, the exact nature and composition of Viking war parties is unknown[5]. One raiding or diplomatic expedition has left direct archaeological traces, at Salme in Estonia, where 41 Swedish males who died violently were buried in two boats accompanied by high-status weaponry[34,35]. Importantly, the Salme boat-burial predates the first textually documented raid (on Lindisfarne, England, in 793) by nearly half a century.

Kinship analysis of the genomes of 34 individuals from the Salme burial reveals four brothers buried side by side and a third degree relative of one of the four brothers (Supplementary Note 4). The Salme group had similar ancestry profiles when compared to the profiles of other Viking burials (Supplementary Notes 10 and 11), suggesting a relatively genetically homogeneous group of people of high status, including close kin.

The five Salme relatives are not the only kin in our dataset. Intriguingly, we also identified two pairs of kin where the related individuals were excavated hundreds of kilometers apart from each other. This dramatically illustrates the mobility of individuals during the VA.

Positive selection in Northern Europe

We looked for SNPs whose allele frequencies changed significantly in the last 10,000 years[36,37] to detect allele frequency shifts in time that cannot be explained by temporal changes in ancestry alone (Supplementary Note 14). Extended Data Figure 8a shows the likelihood ratio scores in favor of selection in the entire 10,000-year period ("general" scan), the period up to 4,000 BP ("ancient" scan) and the period from 4,000 BP up to the present (bottom, "recent" scan).

The strongest candidates for selection are, as expected[38,39], SNPs near the LCT gene, the frequency of which increased after the BA[40,41]. Our dataset traces the frequency of the lactase persistence allele (rs4988235) and its evolution since the BA. Extended Data Figure 8b shows that VA groups had very similar allele frequencies at the LCT lactase persistence SNP to present-day northern European populations. Conversely, BA Scandinavians, and Corded Ware- and Bell Beaker-associated individuals from central Europe, have low frequency despite evidence for milk consumption. Our IA samples have intermediate frequencies, suggesting a rise during this period. The frequency is higher in the BA Baltic Sea region than in BA Scandinavia, consistent with gene flow between the two regions explaining the increasing frequency of lactase persistence in Scandinavia.

Other candidates for selection include previously identified regions—TLR1/TLR6/TLR10, HLA, SLC45A2, and SLC22A4[41]. We also find new candidate regions for selection, with associated trajectories starting before the VA, suggesting shared phenotypes between ancient Vikings and present-day Scandinavians (Supplementary Note 14). These include a region overlapping DCC that is implicated in colorectal cancer[42], and another overlapping AKNA that is involved in the secondary immune response[43].

Evolution of complex traits in Scandinavia

To search for signals of recent population differentiation at SNP markers associated with complex traits, we compared genotypes of VA individuals with those of a present-day Danish panel[44]. We obtained summary statistics from 16 well-powered genome-wide association studies through the GWAS ATLAS[45] and tested for a difference in the distribution of polygenic scores between the two groups (Supplementary note S15). The polygenic scores of VA individuals and present-day Danes differed for three traits: black hair colour ($P = 0.00089$), standing height ($P = 0.019$), and schizophrenia ($P = 0.0096$), though the latter two were not significant after accounting for the number of tests (Extended Data Fig. 7). At the moment, we cannot conclude whether the observed differences in allele frequencies are due to selection acting on these alleles between the VA and the present time or to some other factors (such as more ethnic diversity in the VA sample). A binomial test of the number of black hair colour risk alleles at higher frequency in the VA sample and the present-day sample was also significant (65/41; $P = 0.025$), suggesting the signal is not entirely driven by a few large-effect loci.

Genetic legacy of the Vikings in present-day populations

To test whether present-day Scandinavians share increased ancestry with their respective ancient Viking counterparts, we first computed D-statistics of the form D (YRI, ancient; present-day population 1, present-day population 2), which measure whether an ancient test individual shares more alleles with either present-day population 1 or population 2. Viking Age individuals shift subtly

from Scandinavia towards their present-day counterparts in the distributions of these statistics (Extended Data Fig. 5c; Figs S9.2 and S9.3).

We further examined ancient ancestry in present-day populations using fineSTRUCTURE (Supplementary Note 11, Fig. S11.14). Within Scandinavia, most present-day populations resemble their VA counterparts. The exception is "Swedish-like" ancestry, present at only 15-30% within Sweden, with one Swedish cluster closer to ancient Finnish, and a second more closely related to Danes and Norwegians. "Danish-like" ancestry is now high across the whole region.

Outside of Scandinavia, the genetic legacy of the Vikings is consistent, though limited. A small Scandinavian ancestry component is present in Poland (up to 5%). Within the British Isles, it is difficult to assess how much of the "Danish-like" ancestry is due to pre-existing Anglo-Saxon ancestry, but the VA contribution does not exceed 6% in England (Supplementary Note 11). The genetic impacts are stronger in the other direction. While some "North Atlantic-like" individuals in Orkney became culturally Scandinavian, others found themselves in Iceland, Norway, and beyond, leaving a genetic legacy that persists today. Present-day Norwegians vary between 12 and 25% in "North Atlantic-like" ancestry; this ancestry is more uniformly 10% in Sweden.

Discussion

Our genomic analyses shed light on long-standing questions raised by historical sources and archaeological evidence of the VA. We largely confirm the long-argued movements of Vikings outside Scandinavia: Danish Vikings going to Britain, Norwegian Vikings moving to Ireland, Iceland, and Greenland, and Swedish Vikings sailing east towards the Baltic and beyond. However, we also see ancient "Swedish-like" and FL ancestry in the westernmost fringes of Europe, and "Danish-like" ancestry in the east, defying modern historical groupings. It is likely that many such individuals were from communities with mixed ancestries, thrown together by complex trading, raiding, and settling networks that crossed cultures and the continent.

During the VA, different parts of Scandinavia were not evenly connected, leading to clear genetic structure in the region. Scandinavia likely comprised a limited number of transport zones and maritime enclaves[46] with active external contacts, and limited external gene flow into the rest of the Scandinavian landmass. Some VA Scandinavian locations are relatively homogeneous, particularly mid-Norway, Jutland, and the Atlantic settlements. This contrasts with the strong genetic variation of populous coastal and southern trading communities such as in the islands Gotland and Öland[47–49]. The high genetic heterogeneity in coastal communities implies increased population size, extending both spatially and further back in time the urbanization model for the Late VA city of Sigtuna proposed by Krzewińska et al.[10], who suggested that more cosmopolitan trading centers were already present at the end of the VA in Northern Europe. The formation of large-scale trading and cultural networks that spread people, goods, and warfare took time to affect the heartlands of Scandinavia, which retained pre-existing genetic differences into the medieval period.

Lastly, our findings show that Vikings were not simply a direct continuation of the Scandinavian IA groups. Instead, we observe foreign gene flow from the south and east into Scandinavia, starting in the IA, and continuing throughout the duration of the VA from an increasing number of sources. Many VA individuals have high levels of non-Scandinavian ancestry, both within and outside Scandinavia, suggesting ongoing gene flow across Europe.

Acknowledgements

This work was supported by the Mærsk Foundation, the Lundbeck Foundation, the Novo Nordisk Foundation, the Danish National Research Foundation, University of Copenhagen (KU2016), and the Wellcome Trust (grant nos. WT104125MA). E.W. would like to thank St. John's College, Cambridge for providing an excellent environment for scientific thoughts and collaborations. S.R. was supported by the Novo Nordisk Foundation (NNF14CC0001). F.R. was supported by a Villum Fonden Young Investigator Award (project no. 00025300). G.S. and E.C. were supported by a Marie Skłodowska-Curie Individual Fellowship "PALAEO-ENEO", a project funded by the European Union EU Framework Programme for Research and Innovation Horizon 2020 (Grant Agreement number 751349). R.M. was supported by an EMBO Long-Term Fellowship (ALTF 133-2017). M.C. is supported by the Canada Research Chairs Program (231256), the Canada Foundation for Innovation (36801), and the British Columbia Knowledge Development Fund (962-805808). I.Mo. was supported by a YDUN grant from Independent Research Fund Denmark (DFF-4090-00244) and a Villum Fonden Young Investigator Award (project no. 19114). N.G. was supported by the Program of Fundamental Scientific Research of the State Academies of Sciences, Russian Federation, State Assignment No. 0184-2019-0006. The authors thank the iPSYCH Initiative, funded by the Lundbeck Foundation (grant nos. R102-A9118 and R155-2014-1724), for supplying SNP frequency estimates from the present-day Danish population for comparison with Viking Age samples. We thank Mattias Jakobsson and Anders Götherström for providing preliminary access to the sequencing data of 23 Viking Age samples from Sigtuna. We are also grateful to Marisa Corrente for providing access to the skeletal remains from Cancarro, and Nunzia M. Mangialardi and Marco Maruotti for the useful suggestion; Greenland National Museum and Archives as well as Gotland Museum, for permission to sample their skeletons; John Kavanagh for providing information on his excavation and Laureen Buckley, Denise Keating and Barra Ó Donnabháin for analysing the remains; Richard Breward and Jon Murden from the Dorset County Museum for allowing access to the assemblage for DNA sampling; Carolina Bertilsson, Peter Lingström, Björn Lundberg, Kerstin Lidén and Johanna Andersson for their help in sampling the ancient human remains; Leena Drenzel for permission to sample the skeletons; Catharina Ödman for suggesting the relevant material for this study; Łukasz Stanaszek, Michał Zaitz and the Regional Museum in Cedynia for providing the samples. We thank L. Vinner, A. Seguin-Orlando, K. Magnussen, L. Petersen, C. Mortensen and M.J. Jacobsen at the Danish National Sequencing Centre for producing the analyzed sequences; P.S. Olsen and T. Brand for technical assistance in the laboratories. We thank Richard M. Durbin and James H. Barrett for comments and suggestions. We are grateful to Jim Wilson, Judith Jesch, Erika Harlitz-Kern and Fernando Martín Racimo for their feedback. We also thank the anonymous reviewers for their evaluation and comments.

Contributions

E.W. initiated and led the study.
E.W., A.M., D.J.L., Mar.S., F.R., R.N., K.K., L.H., S.M.S., J.B., N.P., T.W., A.I., M.E.A., M.W.P., N.L., J.A., I.Mo. and A.A. designed the study.

A.M., P.d.B.D., L.C., M.M.B., A.K.F., I.L. and J.S. produced the data.

A.M., D.J.L., Mar.S., F.R., S.R., I.Mo., R.N., T.W., L.C., E.J., A.I., M.W.P., T.K., R.M., G.R., C.B., J.V.M.-M., H.M., A.A., J.C., K.H.I. and M.E.A. analysed or assisted in analysis of data.

E.W., A.M., D.J.L., Mar.S., F.R., S.M.S., K.K., L.H., R.N., M.C., A.I. interpreted the results with considerable input from I.Mo., M.E.A., M.W.P., T.K., H.W., R.M., G.R., T.W., C.H.J., J.A., N.L., N.P., J.B., A.A., M.T.P.G., L.O. and other authors.

E.W., A.M., D.J.L., Mar.S., F.R., S.M.S., K.K., L.H., wrote the manuscript with considerable input from M.C., J.B., N.P., I.Mo., N.L., A.I., R.M., E.J., J.A., M.L.J., C.H.J., M.W.P., M.E.A., G.R. and M.M., with contributions from all authors.

A.M., L.C., M.W.P., H.W., M.M.B., P.d.B.D., A.K.F., M.A., R.A., M.M., E.C., G.S., A.B., A.F., B.Sc., B.Sk., C.A., C.F., D.B., D.P., G.T.-W., H.G., I.L., I.G., I.Ma., I.P., I.M.M., J.M., J.Gi., J.P., J.Gu., L.Si., L.St., L.L., Mae.S., M.F., M.V., M.R., M.B., T.P., M.Sø., N.G., T.C., O.K., O.U., P.F., P.H., S.S., S.A., S.E., V.M., W.B., Y.M., P.P., M.D.J., A.P, D.G.B., M.L.J., J.A., N.L., N.P., M.T.P.G., M.E.A.,J.B. and E.W. excavated, curated, sampled and/or described analysed skeletons; all authors contributed to final interpretation of data.

Competing interests

The authors declare no competing interests.

References

1. Brink, S. & Price, N. *The Viking World*. (Routledge, 2008).
2. Jesch, J. *The Viking Diaspora*. (Routledge, 2015).
3. Eriksen, M. H., Pedersen, U., Rundberget, B. & Axelsen, I. *Viking Worlds: Things, Spaces and Movement*. (Oxbow Books, 2014).
4. Sindbæk, S. M. & Trakadas, A. *The World in the Viking Age*. (Viking Ship Museum in Roskilde, 2014).
5. Sikora, M. *et al.* The population history of northeastern Siberia since the Pleistocene. *Nature* **570**, 182–188 (2019).
6. Lamnidis, T. C. *et al.* Ancient Fennoscandian genomes reveal origin and spread of Siberian ancestry in Europe. *Nat. Commun.* **9**, 5018 (2018).
7. Saag, L. *et al.* The Arrival of Siberian Ancestry Connecting the Eastern Baltic to Uralic Speakers further East. *Curr. Biol.* **29**, 1701–1711.e16 (2019).
8. Hedeager, L. Scandinavia before the Viking Age. in *The Viking World* 35–46 (Routledge, 2008).
9. Hedeager, L. *Iron Age Myth and Materiality: An Archaeology of Scandinavia AD 400-1000*. (Routledge, 2011).
10. Krzewińska, M. *et al.* Genomic and Strontium Isotope Variation Reveal Immigration Patterns in a Viking Age Town. *Curr. Biol.* **28**, 2730–2738.e10 (2018).
11. Lawson, D. J., Hellenthal, G., Myers, S. & Falush, D. Inference of population structure using

dense haplotype data. *PLoS Genet.* **8**, e1002453 (2012).
12. Shepard, D. A Two-dimensional Interpolation Function for Irregularly-spaced Data. in *Proceedings of the 1968 23rd ACM National Conference* 517–524 (ACM, 1968).
13. Hansen, U. L. *Römischer Import im Norden: Warenaustausch zwischen dem Römischen Reich und dem freien Germanien während der Kaiserzeit unter besonderer Berücksichtigung Nordeuropas*. (Det Kongelige nordiske Oldskriftselskab, 1987).
14. Andersson, K. *I skuggan av Rom: romersk kulturpåverkan i Norden*. (Atlantis, 2013).
15. Bill, J. Viking ships and the sea. in *The Viking World* 170–180 (Routledge, 2008).
16. Sindbæk, S. M. The Small World of the Vikings: Networks in Early Medieval Communication and Exchange. *Norwegian Archaeological Review* **40**, 59–74 (2007).
17. Hilberg, V. & Kalmring, S. Viking Age Hedeby and Its Relations with Iceland and the North Atlantic: Communication, Long-distance Trade, and Production. in *Viking Archaeology in Iceland: Mosfell Archaeological Project* 221–245 (Brepols Publishers, 2014).
18. Ebenesersdóttir, S. S. *et al.* Ancient genomes from Iceland reveal the making of a human population. *Science* **360**, 1028–1032 (2018).
19. Helgason, A. *et al.* mtDna and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic ancestry. *Am. J. Hum. Genet.* **68**, 723–737 (2001).
20. Downham, C. Viking ethnicities: a historiographic overview. *History Compass* **10**, 1–12 (2012).
21. Fellows-Jensen, G. Scandinavian place names in the British Isles. in *The Viking World* 391–400 (Routledge London, UK, 2008).
22. Bowden, G. R. *et al.* Excavating past population structures by surname-based sampling: the genetic legacy of the Vikings in northwest England. *Mol. Biol. Evol.* **25**, 301–309 (2008).
23. Leslie, S. *et al.* The fine-scale genetic structure of the British population. *Nature* **519**, 309–314 (2015).
24. Athanasiadis, G. *et al.* Nationwide Genomic Study in Denmark Reveals Remarkable Population Homogeneity. *Genetics* **204**, 711–722 (2016).
25. Loe, L., Boyle, A., Webb, H. & Score, D. *'Given to the Ground': A Viking Age Mass Grave on Ridgeway Hill, Weymouth*. (Dorset Natural History and Archaeological Society, 2014).
26. Wallis, S. *The Oxford Henge and Late Saxon Massacre: With Medieval and Later Occupation at St John's College, Oxford*. (Thames Valley Archaeological Services Limited, 2014).
27. Douglas Price, T., Frei, K. M., Dobat, A. S., Lynnerup, N. & Bennike, P. Who was in Harold Bluetooth's army? Strontium isotope investigation of the cemetery at the Viking Age fortress at Trelleborg, Denmark. *Antiquity* **85**, 476–489 (2011).
28. Price, T. D. & Arneborg, J. The Peopling of the North Atlantic: Isotopic Results from Greenland. *Journal of the North Atlantic* **7**, 164–185 (2014).
29. Arneborg, J. The Norse settlement in Greenland. in *The Viking World* 588–603 (Routledge London, UK, 2008).
30. Dugmore, A. J. *et al.* Cultural adaptation, compounding vulnerabilities and conjunctures in Norse Greenland. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 3658–3663 (2012).
31. Arneborg, J. Norse Greenland: Research into abandonment. in *Medieval Archaeology in Scandinavia and Beyond: History, trends and tomorrow* 247–271 (Aarhus Universitetsforlag, 2015).
32. Lynnerup, N. *The Greenland Norse: a biological-anthropological study*. (Museum Tusculanum Press, 1998).
33. Sindbæk, S. M. Urbanism and exchange in the North Atlantic/Baltic, 600-1000CE. in *Routledge Handbook of Archaeology and Globalization* 553–565 (Routledge, 2016).
34. Peets, J. *et al.* Research results of the Salme ship burials in 2011-2012. *Archaeological*

fieldwork in Estonia **2012**, 43–60 (2012).
35. Douglas Price, T., Peets, J., Allmäe, R., Maldre, L. & Oras, E. Isotopic provenancing of the Salme ship burials in Pre-Viking Age Estonia. *Antiquity* **90**, 1022–1037 (2016).
36. Cheng, J. Y., Racimo, F. & Nielsen, R. Ohana: detecting selection in multiple populations by modelling ancestral admixture components. *bioRxiv* 546408 (2019).
37. Alves, J. M. *et al.* Parallel adaptation of rabbit populations to myxoma virus. *Science* **363**, 1319–1326 (2019).
38. Enattah, N. S. *et al.* Identification of a variant associated with adult-type hypolactasia. *Nat. Genet.* **30**, 233–237 (2002).
39. Bersaglieri, T. *et al.* Genetic signatures of strong recent positive selection at the lactase gene. *Am. J. Hum. Genet.* **74**, 1111–1120 (2004).
40. Allentoft, M. E. *et al.* Population genomics of Bronze Age Eurasia. *Nature* **522**, 167–172 (2015).
41. Mathieson, I. *et al.* Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015).
42. Fearon, E. R. *et al.* Identification of a chromosome 18q gene that is altered in colorectal cancers. *Science* **247**, 49–56 (1990).
43. Siddiqa, A. *et al.* Regulation of CD40 and CD40 ligand by the AT-hook transcription factor AKNA. *Nature* **410**, 383–387 (2001).
44. Pedersen, C. B. *et al.* The iPSYCH2012 case-cohort sample: new directions for unravelling genetic and environmental architectures of severe mental disorders. *Mol. Psychiatry* **23**, 6–14 (2018).
45. Watanabe, K. *et al.* A global view of pleiotropy and genetic architecture in complex traits. *bioRxiv* 500090 (2018).
46. Westerdahl, C. The maritime cultural landscape. *Int. J. Naut. Archaeol.* **21**, 5–14 (1992).
47. Hyenstrand, Å. *Ancient monuments and prehistoric society*. (Central board of national antiquities [Riksantikvarieämbetet], 1979).
48. Callmer, J. Territory and dominion in the Late Iron Age in southern Scandinavia. *Regions and reflections: In honour of Märta Strömberg* 257–273 (1991).
49. Jakobsen, J. G. G. & Dam, P. *Atlas over Danmark: Historisk-Geografisk Atlas*. (Det Kongelige Danske Geografiske Selskab, 2008).

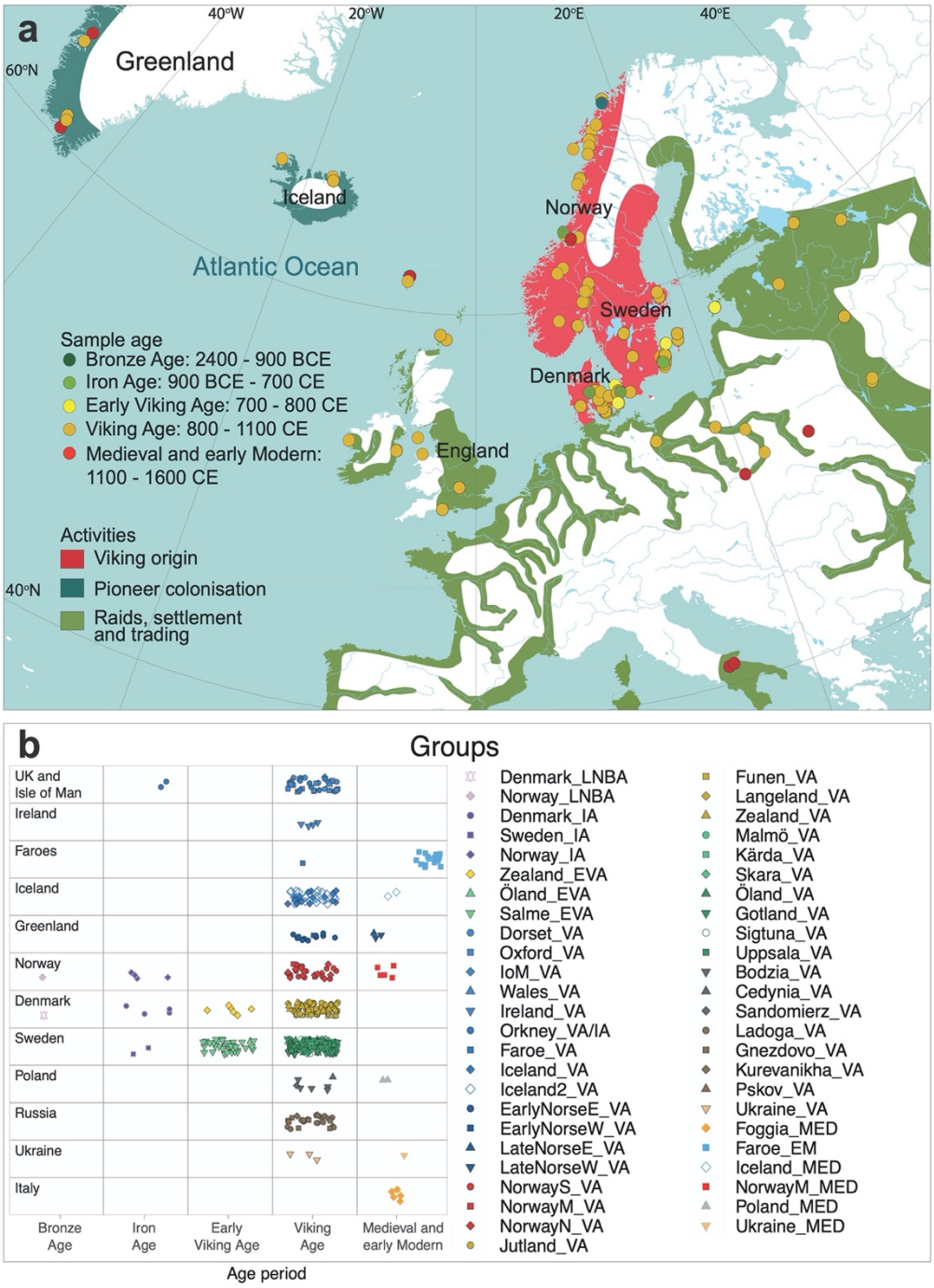


Fig. 1: Viking Age genomic dataset overview. a, Map of the "Viking World" from 8th till 11th centuries, showing geographic location and broad age category (coloured symbols) of sites with new ancient samples reported in this study. **b,** all new ancient individuals from this study (n=442) and published VA samples from Sigtuna[10] and Iceland[18] categorized based on their spatio-temporal origin. The ancient samples are divided into the following five broad categories: Bronze Age (BA), Iron Age (IA), Early Viking Age (EVA), VA and Medieval (MED) / early Modern (EM). Random jitter has been added along the x-axis in each category to aid visualization. LNBA - Late Neolithic/Bronze Age; NorseW - Norse Western settlement; NorseE - Norse Eastern settlement; NorwayS - southern Norway; NorwayN - northern Norway; NorwayM - middle Norway.

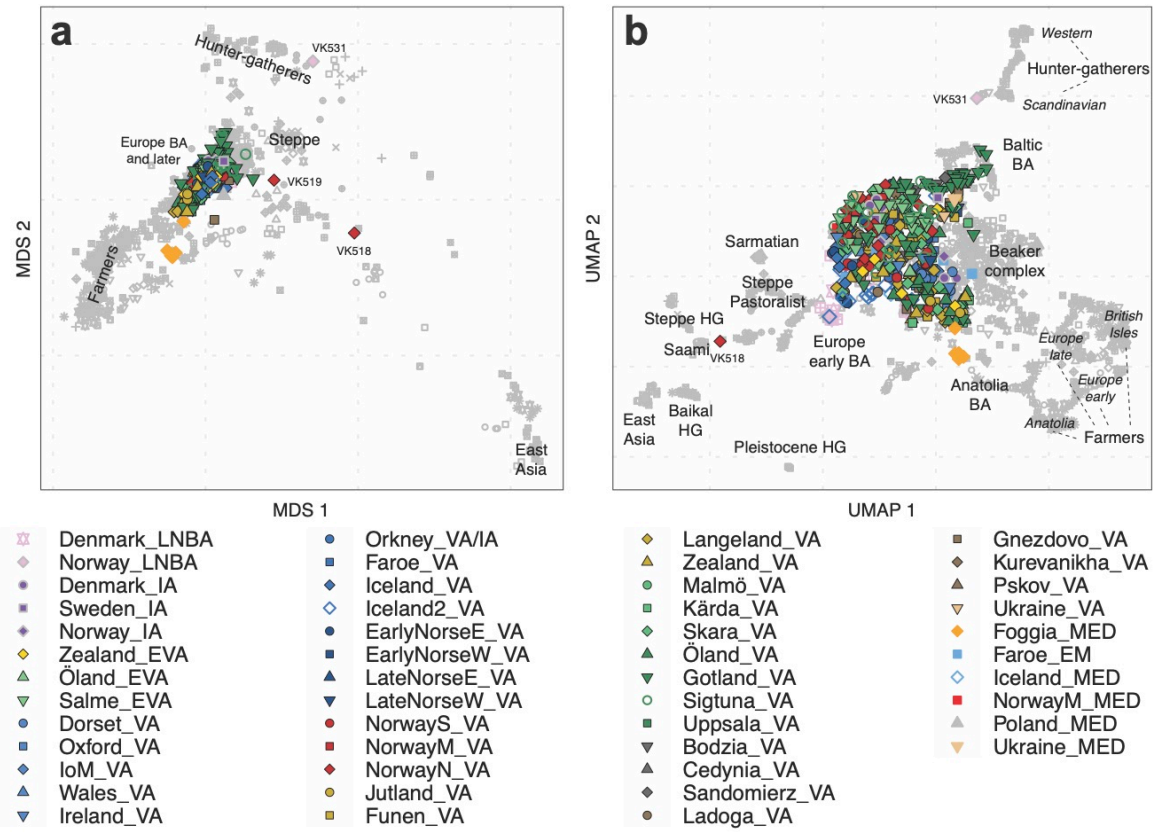


Fig. 2: Genetic structure of VA samples. **a**, Multidimensional scaling (MDS) of n=1,305 ancient genomes, based on a pairwise IBS sharing matrix of the VA and other ancient samples (Supplementary Table 3). Outlier individuals with hunter-gatherer (VK531) or Saami-related ancestry (VK518, VK519) are highlighted. **b**, UMAP analysis of the same dataset as in plot (**a)**, with fine-scale ancestry groups highlighted.

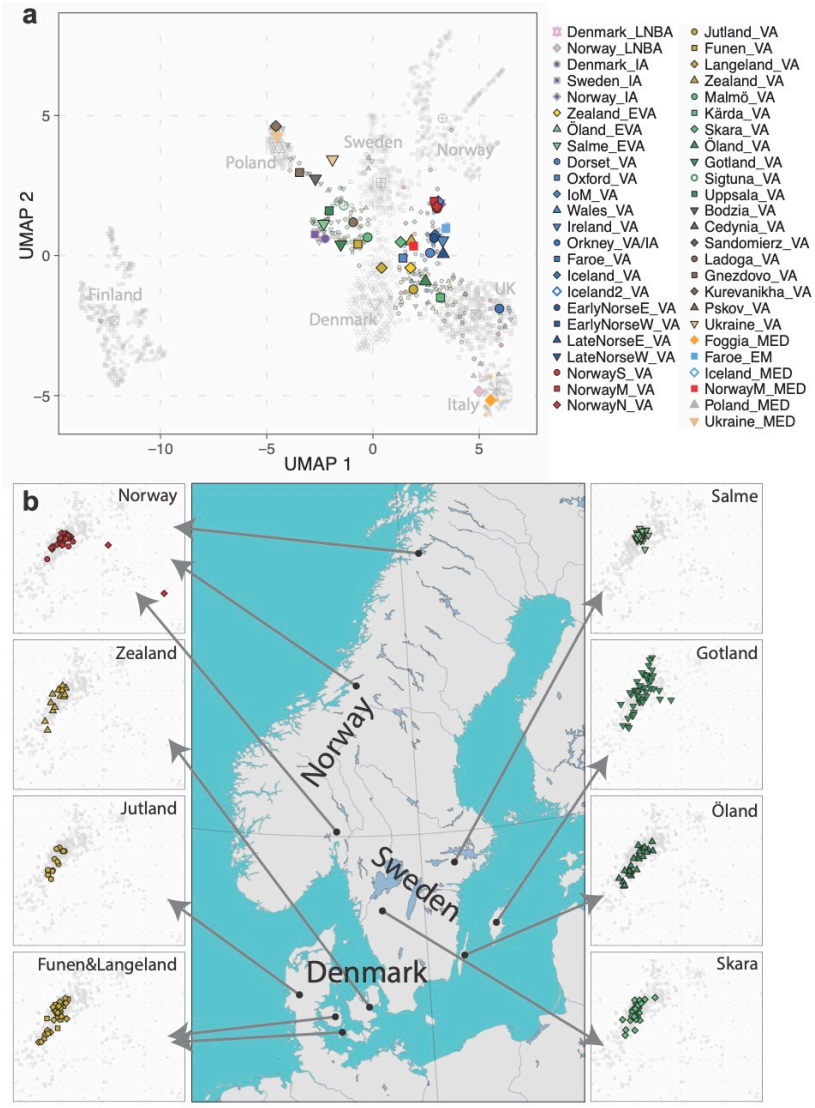


Fig. 3: Genetic structure and diversity of ancient samples. **a**, Uniform manifold approximation and projection (UMAP) analysis of n = 1,624 ancient and modern Scandinavian individuals based on the first 10 dimensions of MDS using IBD segments of imputed individuals. Large symbols indicate median coordinates for each group. **b**, Genetic diversity in major Scandinavian VA populations. Plots next to the map show MDS analysis based on a pairwise IBS sharing matrix. Here "Norway" represents all the sites from Norway. The scale is identical for all the plots.

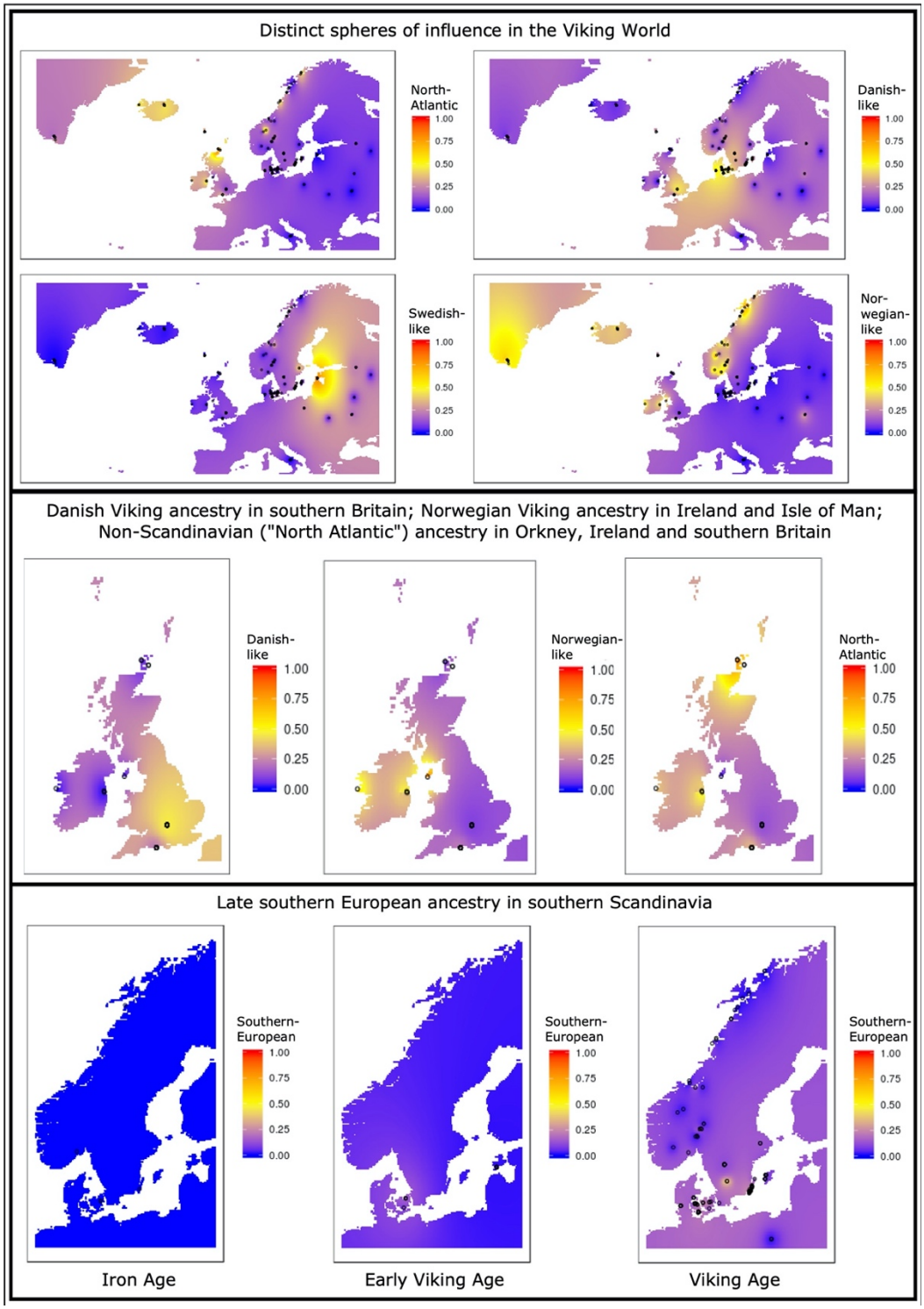


Fig. 4: Spatiotemporal patterns of Viking and non-Viking ancestry in Europe during the IA, EVA and VA. We performed inverse distance weighting interpolation of the ancestry painting proportions of each individual genome on a dense grid of points covering the European continent, to better visualize the distribution of ancestry paintings at different periods (Supplementary Note 12). The "Swedish-like" ancestry is the highest in present-day Estonia due to the ancient samples from the Salme ship burial, which originated from the Mälaren Valley of Sweden, according to archaeological sources. n = 289 genomes used for interpolation.

Methods

Laboratory work

Laboratory work was conducted in the dedicated aDNA clean-room facilities at the Globe Institue, University of Copenhagen according to strict aDNA standards[50,51]. The overwhelming majority of ancient samples were petrous bones and teeth (Supplementary Table 1). The details of DNA extraction can be found in Supplementary Note 2. Double-stranded blunt-end DNA libraries were prepared using Illumina-specific adapters and NEBNext DNA Sample Pre Master Mix Set 2 (E6070) kit. We used Agilent Bioanalyzer 2100 to quantify the amount of the purified DNA libraries. The libraries were sequenced 80 bp single-read chemistry on Illumina HiSeq 2500 machines at the Danish National High-throughput DNA Sequencing Centre.

Bioinformatics analysis and quality assessment

We used AdapterRemoval v2.1.3[52] for removing Illumina adapter sequences keeping only sequences with a minimum length of 30 bp. Adapter-free sequences were mapped against the human reference genome build 37 using BWA v0.7.10 aligner[53] with the seed (-l parameter) disabled for higher sensitivity of ancient DNA reads[54]. DNA sequences were processed with samtools v1.3.1[53] and only sequences with mapping quality ≥30 were kept. Picard v1.127 (http://broadinstitute.github.io/picard) was used to sort the reads and remove duplicates. DNA libraries were combined at sample level and realigned using GATK v3.3.0[55] with Mills and 1000G gold standard indels. At the end, realigned bams had the md-tag updated and extended BAQs calculated using samtools calmd. Read depth and coverage were determined using pysam (http://code.google.com/p/pysam/) and BEDtools[56]. The mapping statistics for the ancient samples are summarized in Supplementary Table 2.
We used mapDamage v2.0 to obtain approximate bayesian estimates of damage parameters[57]. Data authenticity was assessed by estimating the rate of mismatches to the consensus mitochondrial sequence using contamMix[58] and Schmutzi[59] as well as the excess of heterozygous positions in male haploid X chromosomes using ANGSD[60]. The sex of ancient individuals was determined by calculating the Rγ parameter[61].

Uniparental haplogroup determination and kinship analysis

The mitochondrial haplogroups of the ancient individuals were assigned using haplogrep[62]. To get the mtDNA consensus sequences, we aligned the trimmed reads of ancient samples to the human mitochondrial reference genome: revised Cambridge Reference Genome (rCRS). Base quality ≥ 20 and mapping quality ≥ 30 filtering options were applied. Only SNPs at sites ≥ 3X coverage were considered for consensus calling using samtools mpileup/bcftools v1.3.1[53].
We identified male Y chromosome lineages using the pathPhynder workflow (https://github.com/ruidlpm/pathPhynder) and Yleaf v2[63]. For the latter, the analysis was restricted to 26,083 biallelic SNPs from the ISOGG (International Society of Genetic Genealogy) 2019 database (https://isogg.org/tree/ISOGG_YDNA_SNP_Index.html).
We used NgsRelate[64] to detect family relationships between all pairs of individuals. NGSrelate is a maximum-likelihood based program that for a pair of individuals based on genotype likelihoods estimates the three coefficients, k0, k1 and k2, which denote the proportions of the genome where the pair of analyzed individuals share 0, 1 and 2 alleles identical by descent, respectively. We only included the 376 samples with sequencing depth above 0.1X for the analysis. From these we estimated GLs and allele frequencies with ANGSD[60] using the SAMtools GL model (-gl 1) including

reads with MapQ ≥ 30 and bases with baseQ ≥ 20. We only estimated GLs and allele frequencies for the autosomal transversion sites where 1000 Genomes CEU population has a minor allele frequency of 0.05 resulting in 1,752,719 sites. READ[65] was used to confirm the degree of relatedness between pairs of individuals. The pedigree reconstructions based on the kinship coefficients were conducted using PRIMUS - Pedigree Reconstruction and Identification of a Maximum Unrelated Set[66].

Imputation

We imputed the genotypes of 298 ancient samples (289 from this study + 9 from the study by Krzewińska et al.[10]) that had a sequencing depth greater than 0.5X. We used Beagle v4.1[67] for imputations based on the genotype likelihood data, which was first estimated by GATK v3.7.0 UnifiedGenotyper. To generate the genotype data we only called biallelic sites present in the 1000G dataset and only the observed alleles (--genotyping_mode GENOTYPE_GIVEN_ALLELES). The resulting VCF files were filtered by setting genotype likelihoods to 0 for all three genotypes (e.g. hom ref, het and hom alt) for sites with potential deamination (C>T and G>A) as described by Martiniano et al.[68]. Following this, the per-individual vcfs were merged using bcftools-v1.3.1. The combined VCF were then split into 15,000 markers each and imputed separately using beagle-4.0 using the 1000G phase3 map included with beagle (*.phase3.v5a.snps.vcf.gz and plink.chr*.GRCh37.map) with input through the genotype likelihood option. Run time for imputing using beagle was approximately 280,000 core hours.

Merge with existing panels

Scandinavian panel: To assess the genetic relationships of various Viking Age groups with their present-day counterparts we constructed a reference panel enriched with Scandinavian populations based on published datasets: the EGAD00010000632 dataset from Leslie et al.[23] (UK dataset) and the EGAD00000000120 dataset from The International Multiple Sclerosis Genetics Consortium & The Wellcome Trust Case Control Consortium 2[69] (EU dataset), see Supplementary Note 6 for details. Seven most relevant populations from Denmark, Sweden, Norway, Finnland, Poland, UK and Italy were considered (n=1464) with a total number of 414,264 SNPs. The CHB (Han Chinese) and YRI (Yoruba) populations from the 1000 Genomes project phase 3 database were merged to this panel as outgroups.

1000 Genomes panel: We used a set of 1,995 individuals from 20 populations (excluding individuals from the AMR super-population as well as admixed ASW and ACB populations) of the 1000 Genomes project phase 3 release 5 (ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20130502/). We restricted the dataset to a set of 12,731,663 biallelic transversion SNPs located within the 'strict' mappability mask regions (ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20130502/supporting/accessible_genome_masks/).
Analyses of phenotype associated SNPs were carried out using five European-ancestry populations (IBS/Spanish; TSI/Tuscan; CEU/Utah Residents with Northern and Western European Ancestry; GBR/British; FIN/Finnish) along with CHB (Han Chinese) and YRI (Yoruba) as outliers. These were used to assess genome-wide allele frequencies for various SNPs associated with pigmentation phenotypes and lactose intolerance.

Ancient panels: We constructed datasets for population genetic analyses by merging the newly sequenced Viking Age individuals as well as other previously published ancient individuals[40,41,68,70–96] with the two modern reference panels described above. Ancient individuals were represented with "pseudo-haploid" genotypes, obtained by randomly sampling an allele passing filters (mapping quality ≥ 30 and base quality ≥ 30), further requiring that it matched one of the two alleles observed

in the reference panel (Supplementary Table 3). For high coverage ancient and modern individuals, we used diploid genotypes obtained using samtools / bcftools as previously described.

Clustering analyses

Based on the pseudohaploid individuals from the "ancient panels" we ran ADMIXTURE[97] by thinning the dataset for linkage disequilibrium using plink with recommended settings (--indep-pairwise 50 10 0.1). This dataset contained 1324 individuals for 151,235 markers for the autosomal chromosomes. Only samples with >20,000 SNPs overlapping with the "Human Origins panel" were kept in the analysis, resulting in 378 samples from this study. We did 50 replicates with different seeds for k=2 to k=10. We used pong[98] to identify the best run for each K and similar components between different Ks.

The large number of ancient individuals included in the analysis panels facilitates genetic clustering using the ancient individuals themselves, rather than projecting them on axes of variation inferred from modern populations. We carried this out using multi-dimensional scaling (MDS) on a distance matrix obtained from pairwise IBS sharing between individuals, using the 'cmdscale' function in R. We performed the main genetic clustering on a set of 1,306 ancient Eurasian individuals with > 50,000 SNPs with genotype data, restricting to the batch-corrected SNP set described in Supplementary Note 8. Results from the batch-corrected MDS were combined with further dimensionality reduction using uniform manifold approximation and projection (UMAP), implemented in the 'uwot' package in R.

Population genetics

We used f_4 statistics to investigate allele sharing between sets of test individuals and different modern and ancient groups (Supplementary Note 9). To characterize the deep ancestry relationship of the study individuals we calculated f_4 (YRI, Test individual; Barcin_EN.SG, Yamnaya_EBA.SG) for all ancient Europeans from the BA onwards (1000 Genomes panel merge). This statistic contrasts genetic affinities of the test individuals with two major ancestry groups contributing to the gene pool of ancient Europeans from the Bronze Age onwards: Anatolian farmers and Steppe pastoralists. Genetic continuity with Scandinavian Iron Age groups was investigated using f_4 (YRI, Test group; Test individual, Scandinavia IA group) (1000 Genomes panel merge). This statistic measures whether a test individual is consistent with forming a clade with Scandinavian IA groups to the exclusion of a test group from outside of Scandinavia. Genetic affinities between ancient groups and present-day populations were investigated using f_4(YRI, Test individual; Present-day test population, present-day reference population) (Scandinavian panel).

Ancestry modelling using qpAdm

We estimated ancestry proportions of VA groups using *qpAdm*[70], which is based on f_4-statistics of the from f4 (X,O1;O2,O3), where X is either the source or target population, and O1/O2/O3 are triplets of outgroups to the source/target groups. To minimize batch effects and/or biases due to ancient DNA damage or SNP ascertainment, we used a set of 1,800,038 transversion-only sites that were found polymorphic with minor allele frequency ≥ 0.5% and missing genotype rate of ≤ 15% in the 1000 Genomes panel merge.

Genetic diversity

The genetic diversity of ancient groups was assessed using "conditional nucleotide diversity" as previously described[73]. For this analysis, pairwise differences between individuals were calculated using SNPs polymorphic in an outgroup population (YRI) and with a minor allele count ≥ 5 in the 1000 Genomes merge.

IBD analysis

The imputed genotypes of 298 individuals were used to infer genomic segments shared via identity-by-descent (IBD) within the context of a reference panel of 1,464 present-day Europeans, using IBDseq[99] (version r1206) with default parameters. We conducted genetic clustering by MDS on a distance matrix obtained from pairwise IBD sharing and UMAP to reveal fine-scale population structure among Viking Age individuals.

Painting

To assess the fine-scale variation in genetic ancestry proportions of VA individuals we used Chromosome Painting[11]. The following describes the general workflow of the Chromosome Painting analysis, see Supplementary Note 11 for details.

1. Create a modern reference panel using 1675 modern individuals sampled from Northern Europe, using the standard FineSTRUCTURE pipeline:
   - Apply ChromoPainter to paint all modern individuals using the remaining individuals as donors using fs2.0.8. Related individuals were identified through increased haplotype similarity, and admixed individuals were identified by their finestructure clustering. These were removed leading to 1554 unrelated individuals, which were re-painted. Cluster with FineSTRUCTURE, resulting in 40 populations. After removal of small populations and merging of the Chinese (CHB) and African (YRI) sub-populations, this resulted in 23 modern populations with geographical meaning.
   - Call the resulting clustering the "Modern Reference Panel", which consists of 23 Modern Surrogate populations and 23 Modern Donor populations (Figure S11.2).
2. Create an "ancient reference panel" using the modern reference panel:
   - Apply ChromoPainter to paint all ancient individuals using the "Modern Population Palette" (Figure S11.3).
   - Create a supervised "Ancient Population Palette" consisting of 14 populations which either: A: "represent" a modern ancestry direction, or B: are "best associated with" a modern ancestry direction. The paintings consider the average per-individual donor rate to each of the 7 modern populations, normalising each donor label to have mean 1 (Figure S11.4). The individuals that contribute most to a population "represent" it (above a threshold amount chosen by identifying a change-point). The remaining individuals are assigned to the population that they are "best associated with". We create an "Ancient Population Surrogate" for each modern population, consisting of the individuals that "represent" each modern population. For K=7 modern populations, this results in a matrix of K=7 rows (surrogate populations) and 2K=14 columns (donor palette populations) which captures the ancient population structure (Figure S11.6).
3. Infer Ancestry. Learn about population structure in either modern individuals or ancient individuals by painting them with respect to the "ancient population panel" and fitting them as a mixture using the "ancient population surrogates", using the Non-Negative Least Squares (NNLS) implemented in GLOBETROTTER[100] (see Supplementary Section S11) with uncertainty estimated using 100

bootstrap replicates . All samples are analysed by leaving out one individual per donor population so that modern and ancient individuals are exchangeable (as the ancient individual is itself excluded from its own ancient donor population). This is reported in many ways.

- The inferred ancestry results (Supplementary Table S6) are summarized by taking the mean across inferred populations in Figure S11.11, whilst Figure S11.12 shows the means over sample information labels.
- We performed a spatio-temporal regression (Table S11.2) using the model $a_{ik} = \alpha_{jk} t_i + \beta_{jk} x_i + \gamma_{jk} y_i + \varepsilon_{ijk}$, where a_{ik} is the amount of ancestry individual i possesses from population j, t_i is the "age category" of the individual (1=Iron Age, 2=Early Viking Age, 3= Viking Age, 4=Medieval) and (x_i, y_i) are the longitude and latitude of the burial location of the individual.
- The modern ancestry results are estimated using the "Spatial median" instead of the mean, to account for ancestry being constrained in a k-dimensional simplex (Figure S11.14), with uncertainty quantified by bootstrap resampling of individuals (Figure S11.15).

4. Perform sensitivity analyses to ensure that the inference procedure performs as expected. We checked that sequence depth was not associated with cluster membership (Figure S11.7), and that sequence depth did not significantly affect inferred ancestry (Figure 11.8) by downsampling individuals with high depth data available, re-phasing, re-imputing and re-painting them, and assigning ancestry using the above procedure. Results 2X and above were extremely similar, whilst at 1X there was some loss of precision but the broad structure remained clear.
5. Run Principal Components Analysis of the ancient + modern populations painted against our donor populations (Figure S11.9) as well as an all-vs-all ChromoPainter analysis including modern and ancient individuals (Figure S11.10).

Ancestry Diversity Measure

We wish to quantify diversity in ancestry for a population of individuals, with "diverse" meaning a large deviation of individual ancestry estimates from the average ancestry in that population. We compute the average Kullback-Leibler (KL) Divergence for each individual label from the average of that label:

$$D(A^{(l)}) = \frac{1}{n_l} \sum_{i=1}^{n_l} KL(A_i^{(l)} \,||\, p^{(l)})$$

where $A^{(l)}$ is the n_l by K matrix of ancestry estimates in label l, $p^{(l)}$ is the length K vector of average ancestries in that label, and $KL(Q \,||\, P) = \sum_{k=1}^{K} q_k log_2 \left(\frac{q_k}{p_k}\right)$. We performed a simulation study to validate this measure (Supplementary Section S11, Figure S11.13) which allowed us to calibrate the expected diversity as a function of sample size.

Spatiotemporal patterns

To visualise the migration patterns of the Vikings we used inverse distance weighting interpolation implemented in the function 'idw' of the R package *gstat*, to interpolate the proportion of each ancient genome that was attributed by our fineStructure analysis (Supplementary Table 6) to one of the pre-defined ancestry groups: 'UK', 'Denmark', 'Norway', 'Sweden', 'Italy', 'Poland' and 'Finland'. We used the Shepard method of interpolation[12,101] with the weight for a given interpolation location x equal to 1/(d(x,v)^2) where v is the location of an observed sample and d(a,b) is the distance between two points a and b. For plotting maps, we used a Mercator projection and downloaded coastal contours at 1:50m scale from Natural Earth: https://www.naturalearthdata.com/

Lactase persistence and pigmentation SNPs

For ancient populations we estimated the derived 'A' allele frequency of the SNP rs4988235 known to affect expression of the lactase LCT gene. The ancestral "G" allele is responsible for lactase intolerance in adult Europeans[39]. We used ANGSD[60] to estimate the allele frequencies of the ancient population based on the genotype likelihood data. We used the five European populations (CEU – Northern European, FIN – Fins, GBR – British, TSI – Italy, IBS – Spain) and two outgroups (Yoruba – YRI; Chinese – CHB) from the 1000 Genomes Project as comparative groups. We also included the present-day Danish population from the IPSYCH case-cohort study[44] and geographically proximate Iron and Bronze Age populations to trace frequency shifts of SNP rs4988235 through time. We also used ANGSD[60] to estimate the frequencies of 22 SNPs (HIrisPlex[102]) with strongest influence on human pigmentation phenotypes in the VA/EVA Scandinavian population.

Signatures of selection

We aimed to find SNPs whose allele frequencies changed significantly in the last 10,000 years, using our ancient human genomes to look at the frequencies of alleles in the past. We combined our VA and IA genomes with previously published present-day, Bronze Age, Neolithic and Mesolithic sequence data typed at the Human Origins array (see Supplementary Note 6). We filtered for genomes that were younger than 8,000 BCE and that were located within a bounding box encompassing the European continent: 30 < latitude < 75 and -15 < longitude < 45. We then used neoscan in Ohana[36,103] to scan for variants whose allele frequencies were strongly associated with time, after controlling for genome-wide changes in ancestry that might have also occurred over time. We only analyzed sites with a minor allele frequency > 1% (see Supplementary Note 14 for details).

Tracking the evolution of complex traits in Scandinavia

We wanted to examine whether we could identify signals of recent population differentiation of complex traits by comparing genotypes of VA samples excavated in Scandinavia (i.e. Denmark, Sweden and Norway) with those of a present-day Scandinavian population. For the latter, we used imputed genotypes from subjects born in Denmark between 1981-2011 from the IPSYCH case-cohort study[44]. We downloaded summary statistics from the Genome wide association study ATLAS webpage (https://atlas.ctglab.nl)[45], from studies of 16 disease- and anthropometric traits (excluding those related to cognition) published in 2017 or later with SNP heritability estimated at >0.1, sample size of >100,000, and >100 identified genome-wide significant loci. We calculated polygenic risk scores based on independent (R^2<0.1 within 10Mb range) genome-wide significant allelic effects and standardized them to a unit representing the standard deviation of the mean of their distribution. We then removed outliers (anyone with a value for any of the 25 PCs falling more than 4 standard deviations away from the group mean) reiteratively from within each ancestry group (treating the Scandinavian Viking age samples as one ancestry group), and subsequently tested for difference in PRS distribution between Viking age samples and Danish ancestry IPSYCH random population samples using a linear regression model correcting for sex and the 25 principal components.

Data availability

Sequence data are available at the European Nucleotide Archive under accession number PRJEB37976.

Code availability

All raw data are available at the European Nucleotide Archive under accession number PRJEB37976. Functions for calculating f-statistics are available as an R package at GitHub (https://github.com/martinsikora/admixr).

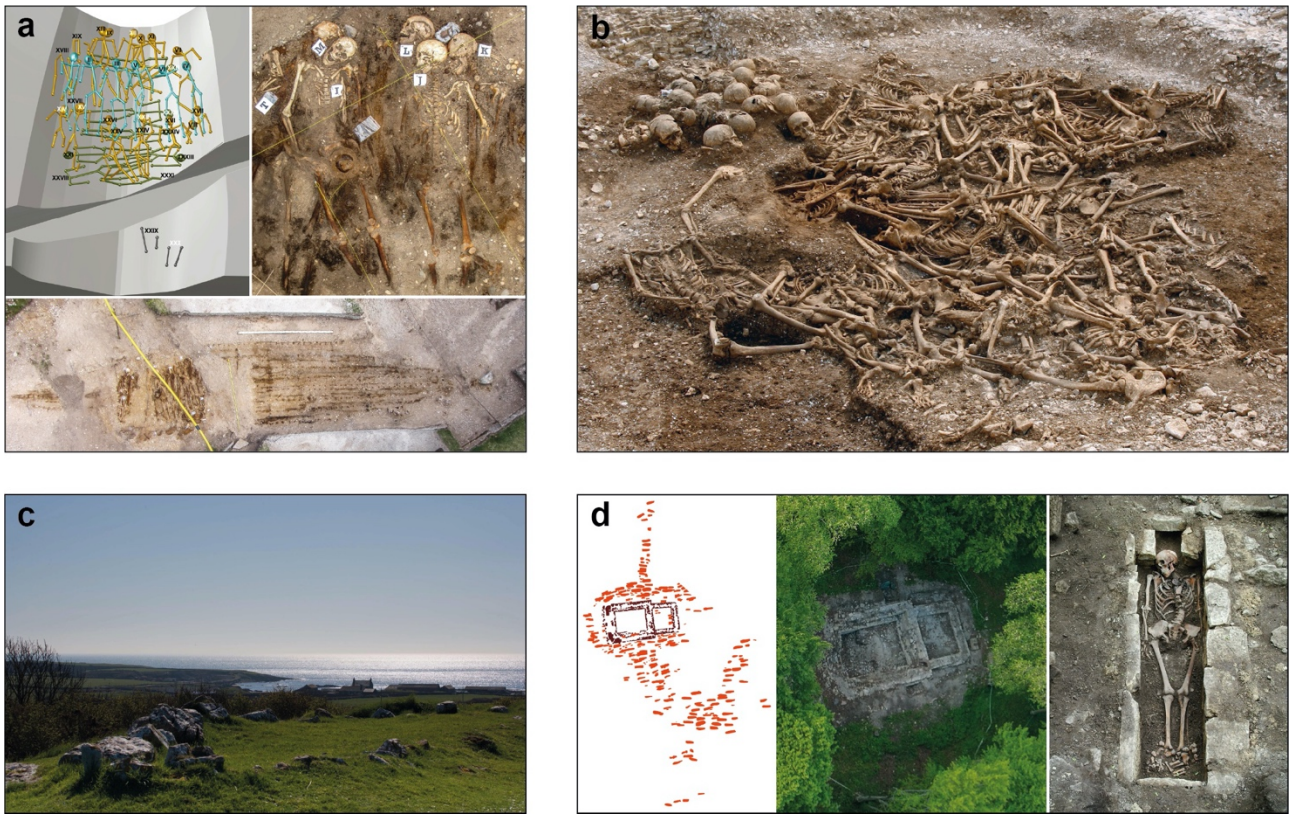
References (Methods section)

50. Willerslev, E. & Cooper, A. Review Paper. Ancient DNA. *Proceedings of the Royal Society B: Biological Sciences* **272**, 3–16 (2005).
51. Gilbert, M. T. P., Bandelt, H.-J., Hofreiter, M. & Barnes, I. Assessing ancient DNA studies. *Trends Ecol. Evol.* **20**, 541–544 (2005).
52. Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **9**, 88 (2016).
53. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
54. Schubert, M. *et al.* Improving ancient DNA read mapping against modern reference genomes. *BMC Genomics* **13**, 178 (2012).
55. DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
56. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
57. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).
58. Fu, Q. *et al.* A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr. Biol.* **23**, 553–559 (2013).
59. Renaud, G., Slon, V., Duggan, A. T. & Kelso, J. Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol.* **16**, 224 (2015).
60. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**, 356 (2014).
61. Skoglund, P., Storå, J., Götherström, A. & Jakobsson, M. Accurate sex identification of ancient human remains using DNA shotgun sequencing. *J. Archaeol. Sci.* **40**, 4477–4482 (2013).
62. Weissensteiner, H. *et al.* HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* **44**, W58–63 (2016).
63. Ralf, A., González, D. M., Zhong, K. & Kayser, M. Yleaf: Software for Human Y-Chromosomal Haplogroup Inference from Next-Generation Sequencing Data. *Mol. Biol. Evol.* **35**, 1291–1294 (2018).
64. Korneliussen, T. S. & Moltke, I. NgsRelate: a software tool for estimating pairwise relatedness from next-generation sequencing data. *Bioinformatics* **31**, 4009–4011 (2015).
65. Monroy Kuhn, J. M., Jakobsson, M. & Günther, T. Estimating genetic kin relationships in

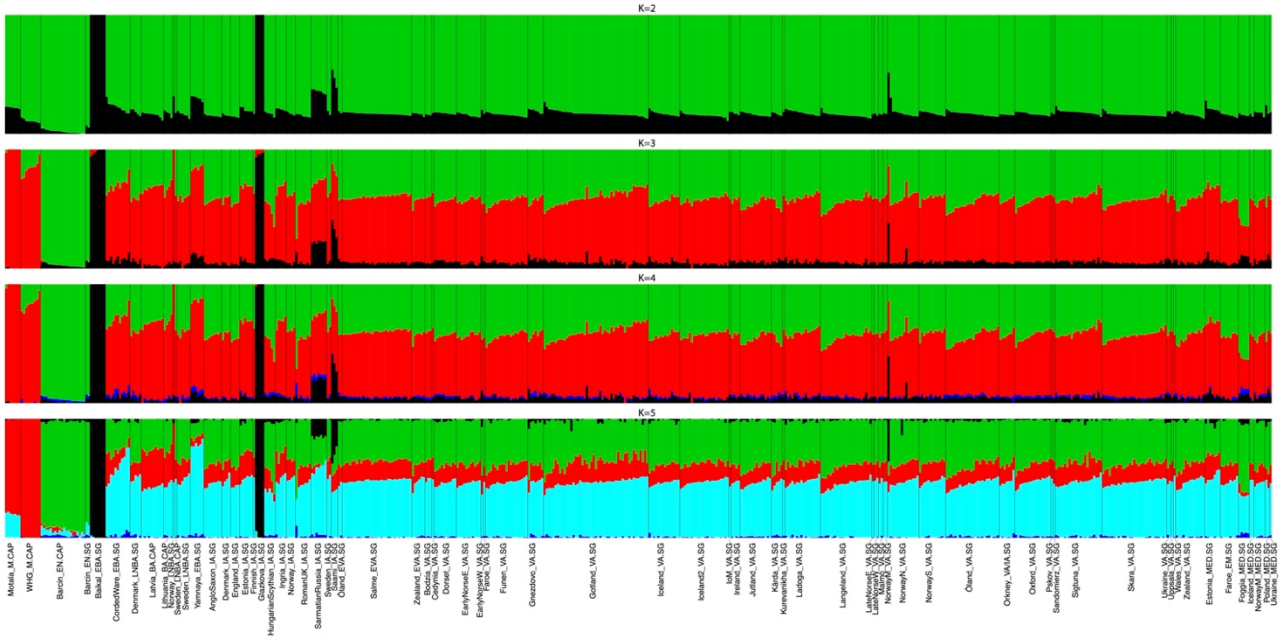
prehistoric populations. *PLoS One* **13**, e0195491 (2018).
66. Staples, J., Nickerson, D. A. & Below, J. E. Utilizing graph theory to select the largest set of unrelated individuals for genetic analysis. *Genet. Epidemiol.* **37**, 136–141 (2013).
67. Browning, S. R. & Browning, B. L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* **81**, 1084–1097 (2007).
68. Martiniano, R. *et al.* The population genomics of archaeological transition in west Iberia: Investigation of ancient substructure using imputation and haplotype-based methods. *PLoS Genet.* **13**, e1006852 (2017).
69. International Multiple Sclerosis Genetics Consortium *et al.* Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* **476**, 214–219 (2011).
70. Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
71. Gamba, C. *et al.* Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* **5**, 5257 (2014).
72. Jones, E. R. *et al.* Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat. Commun.* **6**, 8912 (2015).
73. Skoglund, P. *et al.* Genomic diversity and admixture differs for Stone-Age Scandinavian foragers and farmers. *Science* **344**, 747–750 (2014).
74. Schiffels, S. *et al.* Iron Age and Anglo-Saxon genomes from East England reveal British migration history. *Nat. Commun.* **7**, 10408 (2016).
75. Olalde, I. *et al.* Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* **507**, 225–228 (2014).
76. Sikora, M. *et al.* Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers. *Science* **358**, 659–662 (2017).
77. Fu, Q. *et al.* The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016).
78. Jones, E. R. *et al.* The Neolithic Transition in the Baltic Was Not Driven by Admixture with Early European Farmers. *Curr. Biol.* **27**, 576–582 (2017).
79. Seguin-Orlando, A. *et al.* Paleogenomics. Genomic structure in Europeans dating back at least 36,200 years. *Science* **346**, 1113–1118 (2014).
80. Raghavan, M. *et al.* Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505**, 87–91 (2014).
81. Hofmanová, Z. *et al.* Early farmers from across Europe directly descended from Neolithic Aegeans. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 6886–6891 (2016).
82. de Barros Damgaard, P. *et al.* 137 ancient human genomes from across the Eurasian steppes. *Nature* **557**, 369–374 (2018).
83. Günther, T. *et al.* Population genomics of Mesolithic Scandinavia: Investigating early postglacial migration routes and high-latitude adaptation. *PLoS Biol.* **16**, e2003703 (2018).
84. Mittnik, A. *et al.* The genetic prehistory of the Baltic Sea region. *Nat. Commun.* **9**, 442 (2018).
85. Kılınç, G. M. *et al.* The Demographic Development of the First Farmers in Anatolia. *Curr. Biol.* **26**, 2659–2666 (2016).
86. Lazaridis, I. *et al.* Genetic origins of the Minoans and Mycenaeans. *Nature* **548**, 214–218 (2017).
87. de Barros Damgaard, P. *et al.* The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* **360**, (2018).
88. Valdiosera, C. *et al.* Four millennia of Iberian biomolecular prehistory illustrate the impact of prehistoric migrations at the far end of Eurasia. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 3428–3433 (2018).

89. Martiniano, R. *et al.* Genomic signals of migration and continuity in Britain before the Anglo-Saxons. *Nat. Commun.* **7**, 10326 (2016).
90. Mathieson, I. *et al.* The genomic history of southeastern Europe. *Nature* **555**, 197–203 (2018).
91. Gallego Llorente, M. *et al.* Ancient Ethiopian genome reveals extensive Eurasian admixture throughout the African continent. *Science* **350**, 820–822 (2015).
92. Broushaki, F. *et al.* Early Neolithic genomes from the eastern Fertile Crescent. *Science* **353**, 499–503 (2016).
93. Veeramah, K. R. *et al.* Population genomic analysis of elongated skulls reveals extensive female-biased immigration in Early Medieval Bavaria. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 3494–3499 (2018).
94. Amorim, C. E. G. *et al.* Understanding 6th-century barbarian social organization and migration through paleogenomics. *Nat. Commun.* **9**, 3547 (2018).
95. Olalde, I. *et al.* A Common Genetic Origin for Early Farmers from Mediterranean Cardial and Central European LBK Cultures. *Mol. Biol. Evol.* **32**, 3132–3142 (2015).
96. Olalde, I. *et al.* The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature* **555**, 190–196 (2018).
97. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
98. Behr, A. A., Liu, K. Z., Liu-Fang, G., Nakka, P. & Ramachandran, S. pong: fast analysis and visualization of latent clusters in population genetic data. *Bioinformatics* **32**, 2817–2823 (2016).
99. Browning, B. L. & Browning, S. R. Detecting identity by descent and estimating genotype error rates in sequence data. *Am. J. Hum. Genet.* **93**, 840–851 (2013).
100. Hellenthal, G. *et al.* A genetic atlas of human admixture history. *Science* **343**, 747–751 (2014).
101. Fortin, M.-J. & Dale, M. R. T. *Spatial analysis: a guide for ecologists*. (Cambridge University Press, 2005).
102. Walsh, S. *et al.* The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. *Forensic Sci. Int. Genet.* **7**, 98–115 (2013).
103. Cheng, J. Y., Mailund, T. & Nielsen, R. Fast admixture analysis and population tree estimation for SNP and NGS data. *Bioinformatics* **33**, 2148–2155 (2017).

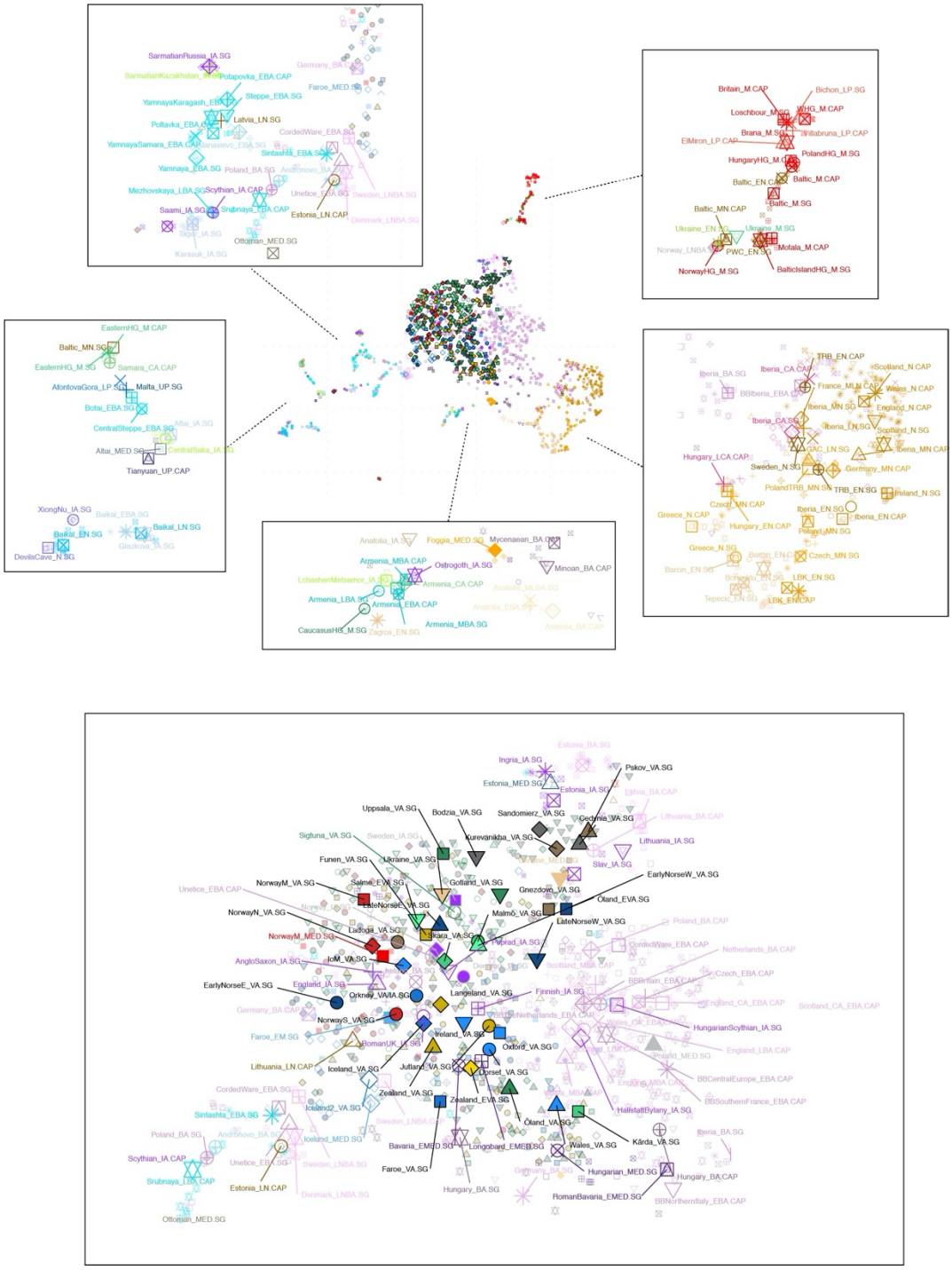
Extended Data Figures



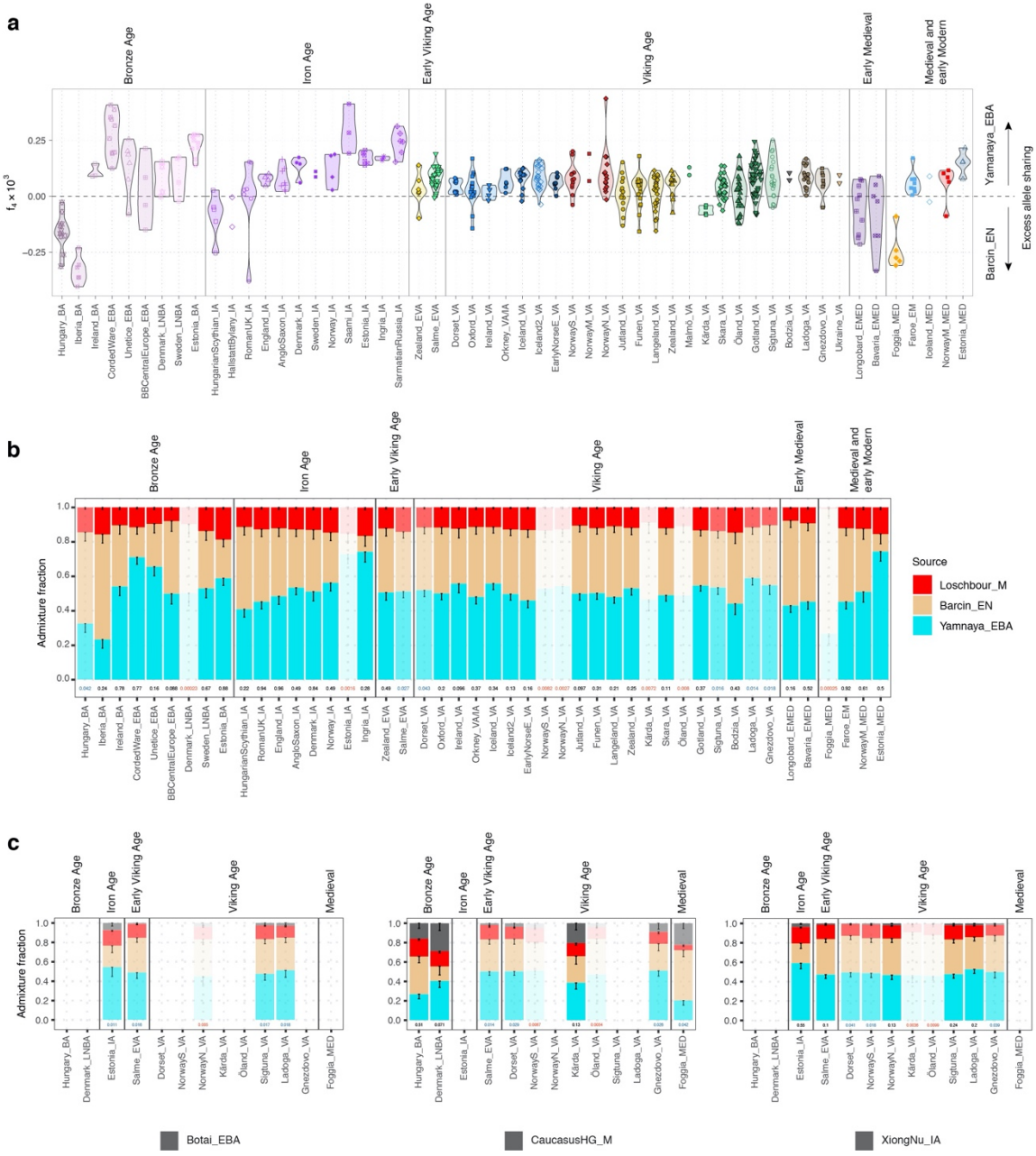
Extended Data Fig. 1: Viking Age archaeological sites
Examples of a few archaeological Viking Age sites and samples used in this study. a, Salme II ship burial site of Early Viking Age excavated in present-day Estonia: schematic representation of skeletons (upper left-hand corner image) and aerial images of skeletons (upper right-hand corner and lower images). b, Ridgeway Hill mass grave dated to the 10th or 11th century, located on the crest of Ridgeway Hill, near Weymouth, on the South coast of England. Around 50 predominantly young adult male individuals were excavated. c, The site of Balladoole: around AD 900, a Viking was buried in an oak ship at Balladoole, Arbory in the south east of the Isle of Man. d, Viking Age archaeological site in Varnhem, in Skara municipality, Sweden: Schematic map of the church foundation (left) and the excavated graves (red markings) at the early Christian cemetery in Varnhem; foundations of the Viking Age stone church in Varnhem (middle) and the remains of a 182 cm long male individual (no. 17) buried in a lime stone coffin close to the church foundations (right).



Extended Data Fig. 2: Model-based clustering analysis
Admixture plot (K=2 to K=5) for 567 ancient individuals spanning 71 different populations. This figure is a subset of most relevant individuals and populations from Figure S7.2, see Supplementary Note 7 for details. This plot consists of 378 ancient samples from this study; VA samples from Sigtuna, Sweden[10] (n=21); Iceland[18] (n=22) and other ancient comparative groups (n=146).

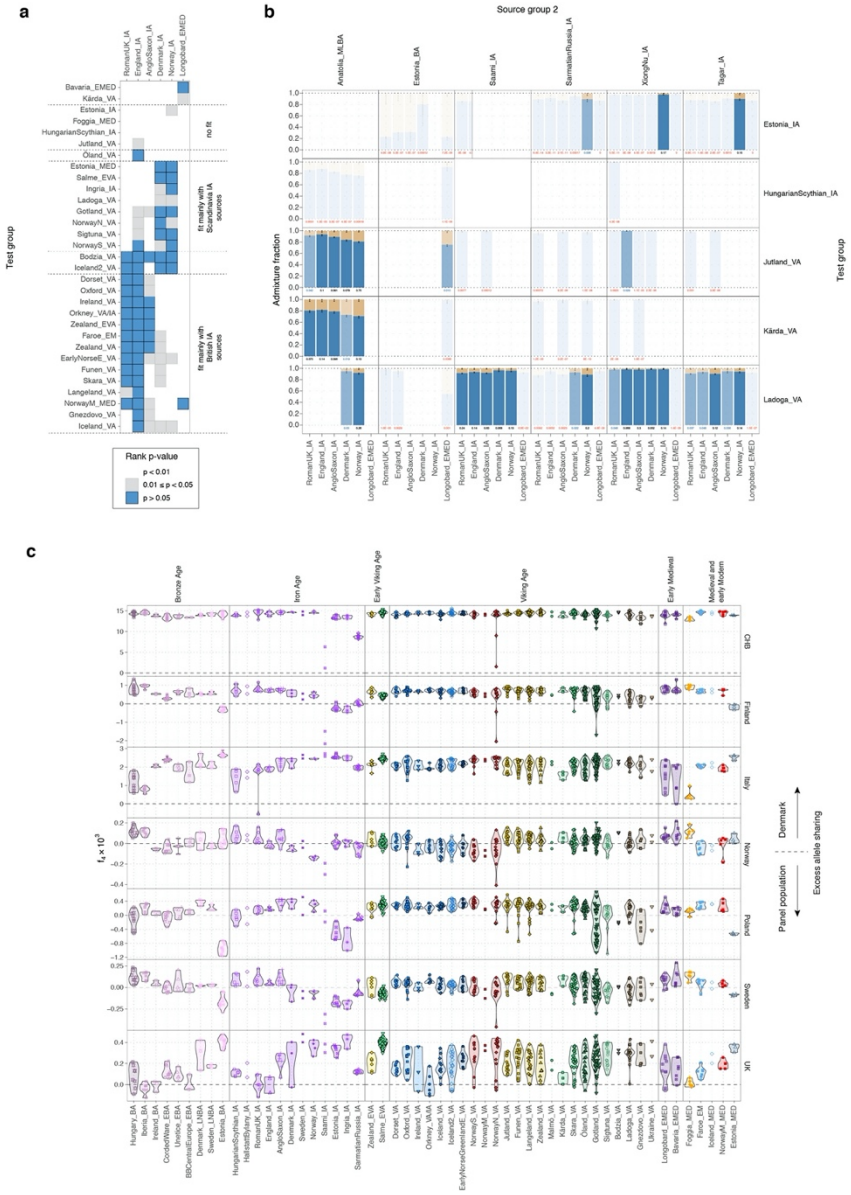


Extended Data Fig. 3: Fine-scale population structure
The point cloud at the top center shows an alternative view of the UMAP result from Figure 2b, with all ancient individuals colored based on analysis group. The framed panels surrounding the point cloud highlight particular ancestry clusters as indicated, with labels and larger symbols corresponding to the median coordinates for the respective group. The larger bottom panel similarly shows median group coordinates for the large central point cloud, which includes the vast majority of European individuals from the Bronze Age onwards.



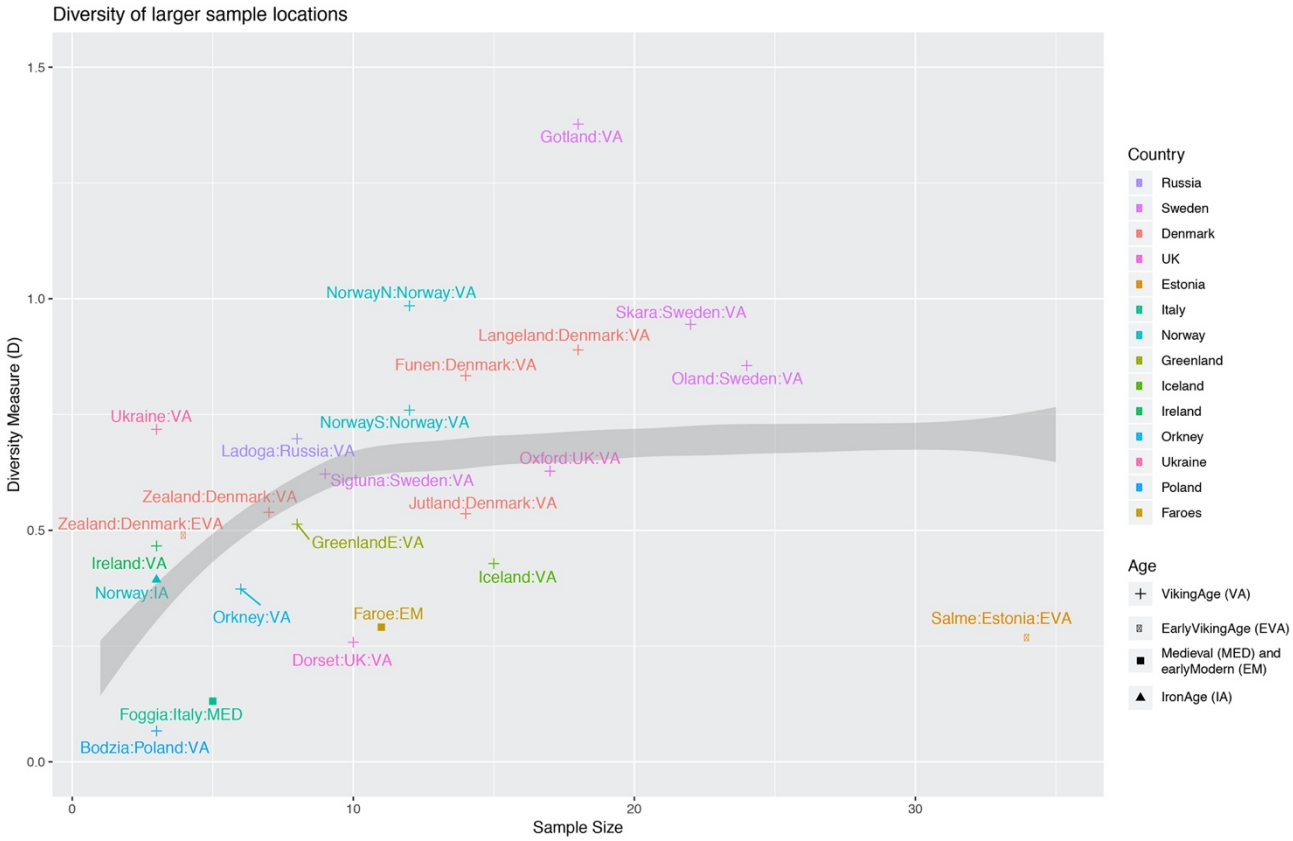
Extended Data Fig. 4: Ancestry modelling for distal sources

a, Contrasting allele sharing between Anatolian farmers (Barcin_EN) and Steppe pastoralists (Yamnaya_EBA) for European individuals from the Bronze Age and later. Violin plots showing distributions of statistics f_4(YRI,test individual;Barcin_EN,Yamnaya_EBA) for n=515 individuals with a minimum of 1,000,000 SNPs with genotypes and groups with at least two such individuals. **b,** Ancestry proportions of analysis groups from the Bronze Age and later inferred using *qpAdm*. Target groups were modelled using three distal sources representing European hunter-gatherer (Loschbour_M), Anatolian farmer (Barcin_EN) and Steppe pastoralist (Yamnaya_EBA) ancestry. Sample sizes for target groups can be found in Supplementary table 10. Error bars indicate standard error obtained from *qpAdm*. **c,** Ancestry proportions of analysis groups for which the three source model was rejected using *qpAdm* ($p < 0.05$). Target groups were modelled including one additional distal source representing either Steppe hunter-gatherer (Botai_EBA), Caucasus hunter-gatherer (CaucasusHG_M) or East Asian-related (XiongNu_IA) ancestry.

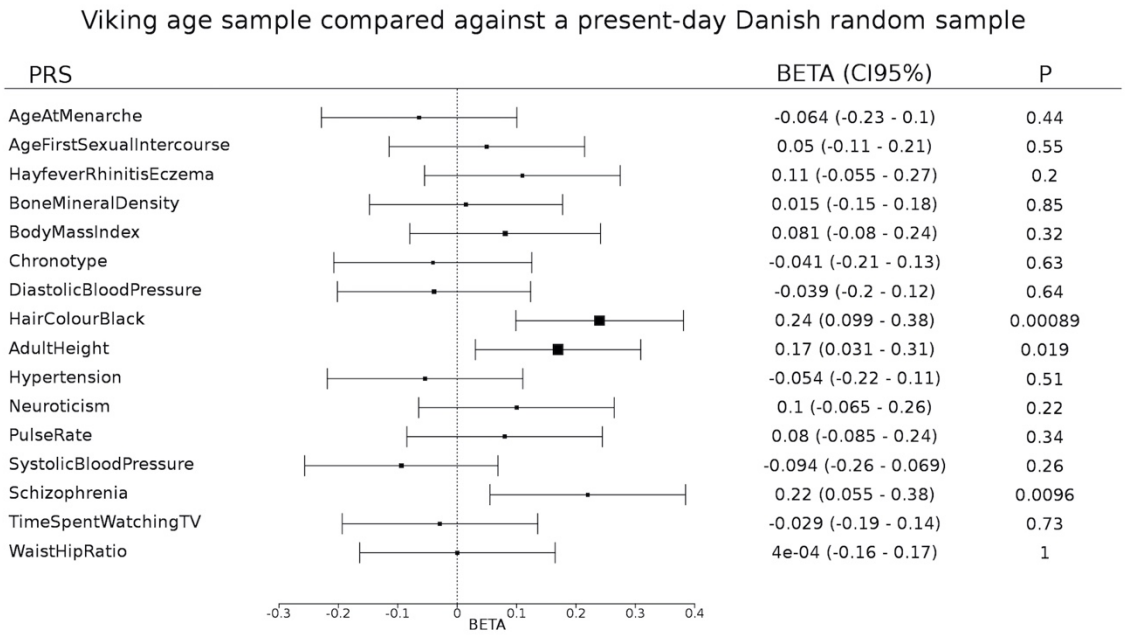


Extended Data Fig. 5: Ancestry modelling for proximate sources

a, Testing for continuity between European Iron Age and later Viking Age and Medieval groups. Couloured squares depict whether a particular target group (row) can be modelled using a single source group (column). P-values for f_4 rank of 0 (corresponding to a single source group) were obtained using *qpAdm* with a set of 15 outgroups which included European Bronze Age groups preceding the source groups. Sample sizes for target groups can be found in Supplementary table 12 **b,** Two-way admixture ancestry proportions of target groups for which a single source was rejected ($p \leq 0.05$). Target groups were modelled using additional proximate Bronze and Iron Age sources. Sample sizes for target groups can be found in Supplementary table 13. For both **a, b,** only ancient groups containing at least three individuals with a minimum of 1,000,000 SNPs with genotypes are plotted **c,** Contrasting allele sharing between present-day Denmark and other populations. Violin plots showing distributions of statistics f_4(YRI,test individual;Panel population,Denmark) for n=489 individuals with a minimum of 50,000 SNPs with genotypes and groups with at least two such individuals. Median values for distributions are indicated with horizontal lines.

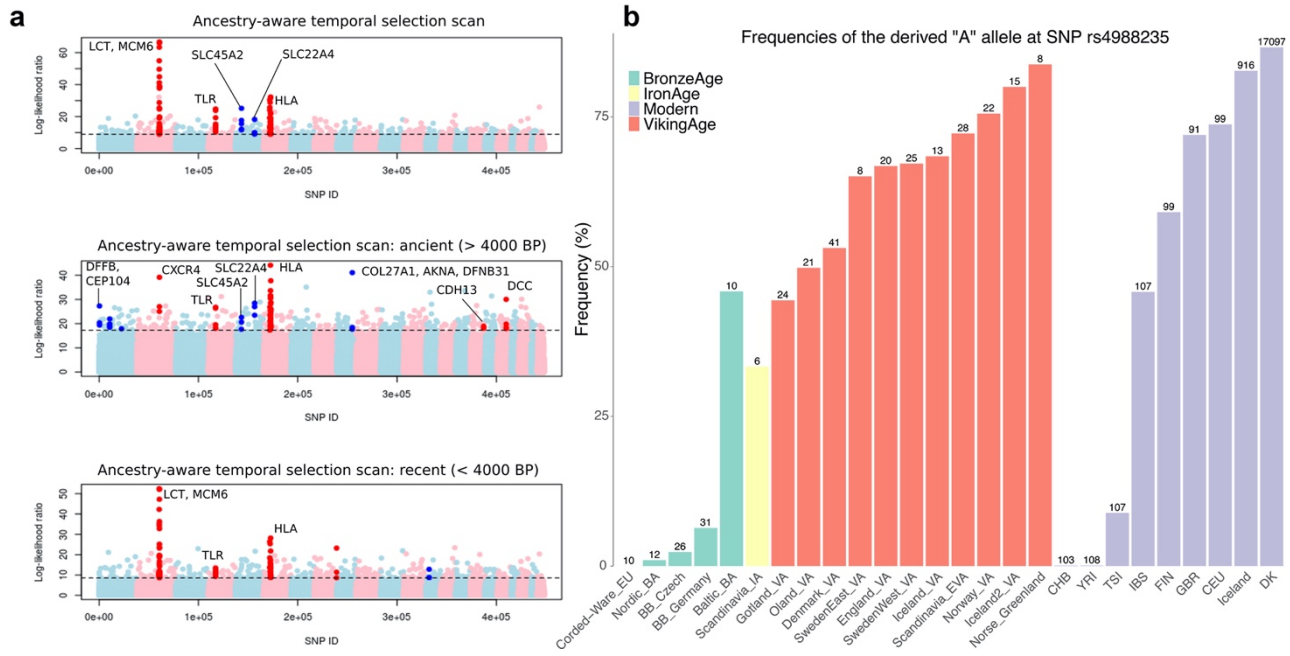


Extended Data Fig. 6: Ancestry diversity of different population groups
Diversity of different labels (i.e. sample locations combined with historical age) are shown as a function of their sample size. The Diversity measure is the Kullback-Leibler divergence from the label means, capturing the diversity of a group with respect to the average of that group; see text for details. Larger values are more diverse, though a dependence on sample size is expected. The simulation expectation for the best-fit to the data (0=0.2) is shown.



Extended Data Fig. 7: Polygenic risk scores

Polygenic risk scores (PRS) for 16 complex human traits in 148 Viking Age samples from Denmark, Sweden and Norway compared against a reference sample of 20,551 Danish-ancestry individuals randomly drawn from all individuals born in Denmark in 1981-2005. The PRS is in each case based on allelic effects for >100 independent genome-wide significant SNPs from recent GWAS of the respective traits and standardised to a mean of 0 and standard deviation of 1 in the entire sample. Difference in PRS was estimated in a linear regression correcting for sex and 25 principal components of overall genetic structure. The plotted BETA indicates the coefficient for the testgroup (Viking Age sample) PRS compared to that of the Danish comparison sample, with error bars indicating the 95% confidence interval of BETA, and P indicating the two-tailed p-value of the corresponding T-test (not corrected for number of tests). Only PRS for black hair color is significantly different between the groups after taking account of multiple testing.



Extended Data Fig. 8: Positive selection in Europe

a, Manhattan plots of the likelihood ratio scores in favor of selection looking at the entire 10,000-year period (top, “general” scan), the period up to 4,000 BP (middle, “ancient” scan) and the period from 4,000 BP up to the present (bottom, “recent” scan). The highlighted SNPs have a score larger than the 99.9% quantile of the empirical distribution of log-likelihood ratios, and have at least two neighboring SNPs (+/- 500kb) with a score larger than the same quantile. n = 1,185 genomes are used in the selection scan. **b**, Frequencies of the derived “A” allele rs4988235 SNP responsible for lactase persistence in humans for different Viking-Age groups, present-day populations from the 1000 Genomes Project as well as relevant Bronze Age population panels. The numbers at the top of the bars denote the sample size on which the allele frequency estimates are based.